

ANALYSIS OF MYNTRA APPAREL

INTRODUCTION

Myntra is an Indian fashion e-commerce company. The company was founded in 2007-2008 to sell personalized gift items. Myntra is a one-stop shop for all your fashion and lifestyle needs. Being India's largest e-commerce store for fashion and lifestyle products, Myntra aims to provide a hassle-free and enjoyable shopping experience to shoppers across the country with the widest range of brands and products on its portal. Myntra understands its shoppers' needs and caters to them with a choice of apparel, accessories, cosmetics, and footwear from over 500 leading Indian and international brands. Prominent brands include Adidas, Nike, Puma, and so many Indian and international brands. The company is headquartered in Bengaluru, Karnataka, and services over 19,000 pin codes across India.





PROJECT QUESTIONS

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."

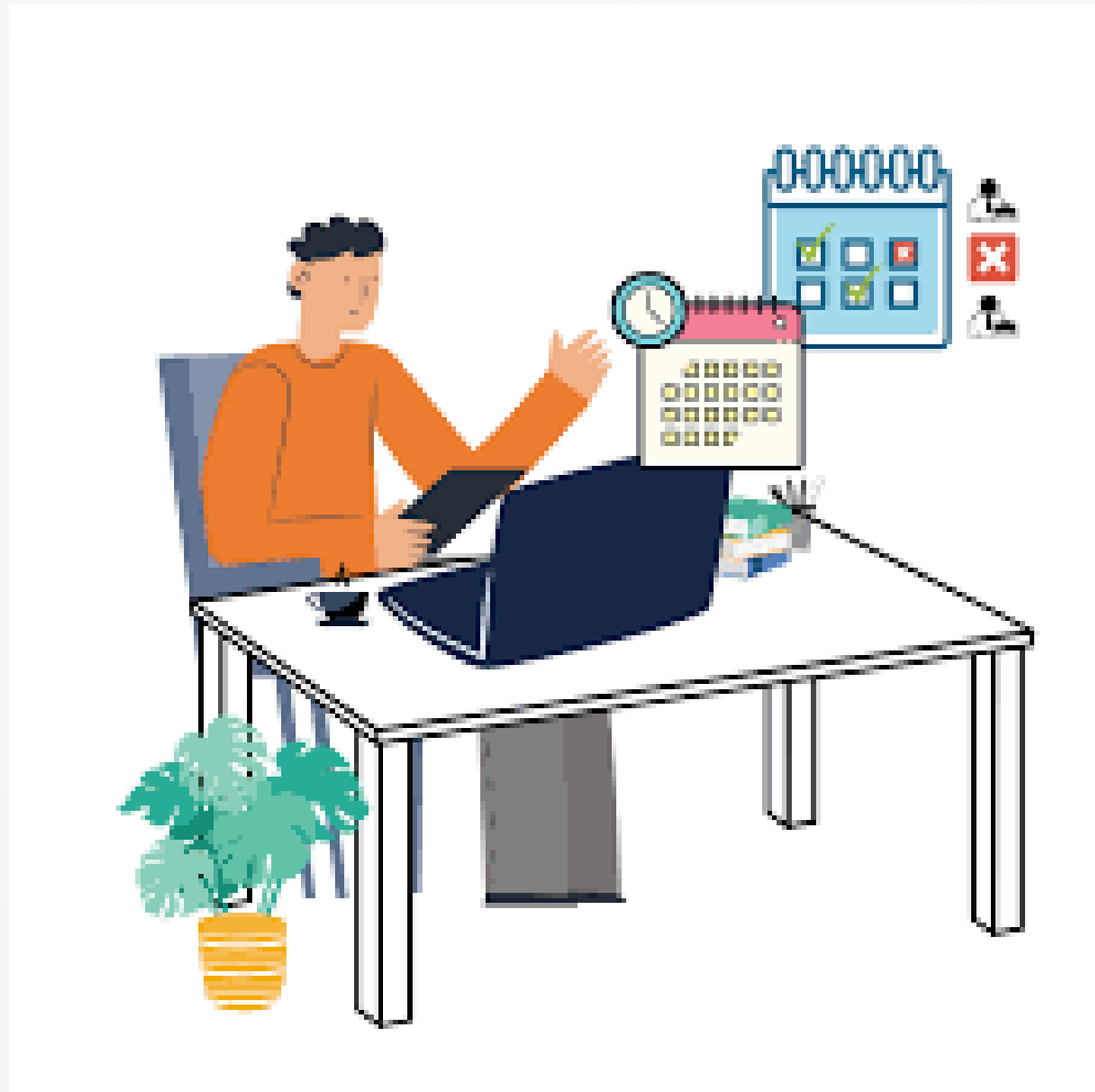
B. Data Analysis

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



Check for duplicate values in your dataset and remove them.

URL	Product_id	BrandName	Category	Individual_category	Description	DiscountPercentage	OriginalPrice	DiscountedPrice	SizeOption	Ratings	Feedback
https://www.roadster.com/jeans/men/roadster-men-jeans-2296012	2296012	Roadster	Bottom Wear	Men	roadster men jeans	824	1499	45% OFF	28, 30, 32,	3.9	
https://www.locomotivestore.com/track-pant/men/locomotive-men-track-pant-13780156	13780156	LOCOMOTIVE	Bottom Wear	Men	locomotive men track-pant	517	1149	55% OFF	S, M, L, XL	4	
https://www.roadster.com/topwear/men/roadster-men-shirts-11895958	11895958	Roadster	Topwear	Men	roadster men shirts	629	1399	55% OFF	38, 40, 42,	4.3	
https://www.zivame.com/women/zivame-women-lingerie-&-shapewear-4335679	4335679	Zivame	Lingerie & Shapewear	Women	zivame women lingerie	893	1295	31% OFF	S, M, L, XL	4.2	
https://www.roadster.com/vest/men/roadster-men-vest-11690882	11690882	Roadster	Vest	Men	roadster men vest	599	999	35% OFF	XS, S, M, L,	4.2	
https://www.mastandhatter.com/hat/men/mast-and-hatter-men-hat-2490950	2490950	Mast & Hatter	Hat	Men	mast and hatter men hat	599	999	40% OFF	XS, S, M, L,	4.4	
https://www.hIGHLANDER.com/track-pant/men/highlander-men-track-pant-6744434	6744434	HIGHLANDER	Bottom Wear	Men	highlander men track-pant	1499	499	60% OFF	30, 32, 34,	3.9	
https://www.mayra.com/vest/men/mayra-men-vest-8439415	8439415	Mayra	Vest	Men	mayra men vest	1395	2699	58% OFF	S, M, L, XL	3.7	
https://www.roadster.com/vest/men/roadster-men-vest-17381394	17381394	Roadster	Vest	Men	roadster men vest	1098	1098		XS, S, M, L,	4.3	
https://www.hereandnow.com/track-pant/men/here-and-now-men-track-pant-2359257	2359257	HERE&NOW	Bottom Wear	Men	here and now men track-pant	2749	2749		28, 30, 32,	3.5	
https://www.hrxbyhri.com/track-pant/men/hrx-by-hri-men-track-pant-7695793	7695793	HRX by Hri	Bottom Wear	Men	hrx by hri men track-pant	2699	2699	55% OFF	S, M, L, XL	4.4	
https://www.roadster.com/track-pant/men/roadster-men-track-pant-10307375	10307375	Roadster	Bottom Wear	Men	roadster men track-pant	699	699		XS, S, M, L,	4.1	
https://www.anubhuteelife.com/track-pant/men/anubhutee-life-men-track-pant-12873874	12873874	Anubhutee Life	Bottom Wear	Men	anubhutee life men track-pant	3399	3399	70% OFF	S, M, L, XL	4.2	
https://www.athena.com/vest/men/athena-men-vest-11634538	11634538	Athena	Vest	Men	athena men vest	2499	2499	50% OFF	XS, S, M, L,	4.3	
https://www.roadster.com/vest/men/roadster-men-vest-2312181	2312181	Roadster	Vest	Men	roadster men vest	799	799	60% OFF	XS, S, M, L,	4	
https://www.hIGHLANDER.com/track-pant/men/highlander-men-track-pant-13842386	13842386	HIGHLANDER	Bottom Wear	Men	highlander men track-pant	1099	1099	53% OFF	39, 40, 42,	4.2	
https://www.vishudh.com/track-pant/men/vishudh-men-track-pant-10473520	10473520	Vishudh	Bottom Wear	Men	vishudh men track-pant	1699	1699	59% OFF	S, M, L, XL	4.2	
https://www.sangria.com/vest/men/sangria-men-vest-10561392	10561392	Sangria	Vest	Men	sangria men vest	3999	3999		XS, S, M, L,	4.3	
https://www.tokyo-talkies.com/track-pant/men/tokyo-talkies-men-track-pant-12391750	12391750	Tokyo Talkies	Bottom Wear	Men	tokyo talkies men track-pant	2149	2149	61% OFF	28, 30, 32,	4.1	
https://www.dressberry.com/track-pant/men/dressberry-men-track-pant-2522986	2522986	DressBerry	Bottom Wear	Men	dressberry men track-pant	999	999	75% OFF	XS, S, M, L,	4.3	
https://www.roadster.com/track-pant/men/roadster-men-track-pant-17385142	17385142	Roadster	Bottom Wear	Men	roadster men track-pant	1298	1298		XS, S, M, L,	4.4	
https://www.anouk.com/track-pant/men/anouk-men-track-pant-12153330	12153330	Anouk	Bottom Wear	Men	anouk men track-pant	719	1599	55% OFF	XS, S, M, L,	3.8	
https://www.enamor.com/track-pant/men/enamor-men-track-pant-1864573	1864573	Enamor	Bottom Wear	Men	enamor men track-pant	849	849		S, M, L, XL	4.2	

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

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

ne wol	893	1295	31% OFF	S, M, L, XL,	4.2	999
ster women white	599	599	35% OFF	XS, S, M, L,	4.2	999
harbour women	599	599	40% OFF	XS, S, M, L,	4.4	999
ander	599	1499	60% OFF	30, 32, 34,	3.9	998
a pink embroider	1395	1395	58% OFF	S, M, L, XL	3.7	998
ster women pack	1098	1098		XS, S, M, L,	4.3	997
now men blue slir	2749	2749		28, 30, 32,	3.5	996
y hrith	1214	2699	55% OFF	S, M, L, XL,	4.2	995
ster men navy blu	699	699		XS, S, M, L,	4.2	995
hutee	1019	3399	70% OFF	S, M, L, XL,	4.3	995
na women black s	2499	2499	50% OFF	XS, S, M, L,	4.3	995
ster women marc	799	799	60% OFF	XS, S, M, L,	4.3	995
ander	516	1099	53% OFF	30, 32, 34,	3.9	998
dh wc	696	1699	59% OFF	S, M, L, XL,	4.2	995
ria women green	3999	3999		XS, S, M, L,	4.3	995
o talkies women o	2149	2149	61% OFF	28, 30, 32,	4.1	995
sberry women lov	999	999	75% OFF	XS, S, M, L,	4.3	995
ster women pack	1298	1298		XS, S, M, L,	4.4	994
ik won	719	1599	55% OFF	XS, S, M, L,	3.8	993
nor black non wir	849	849		S, M, L, XL	4.2	993

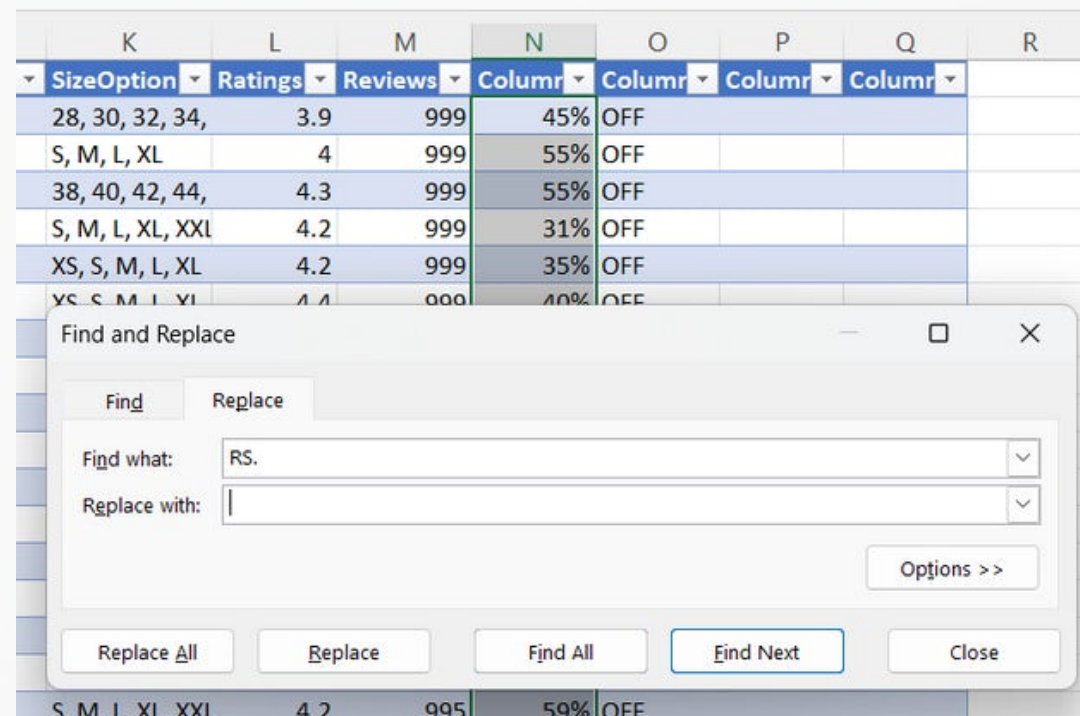
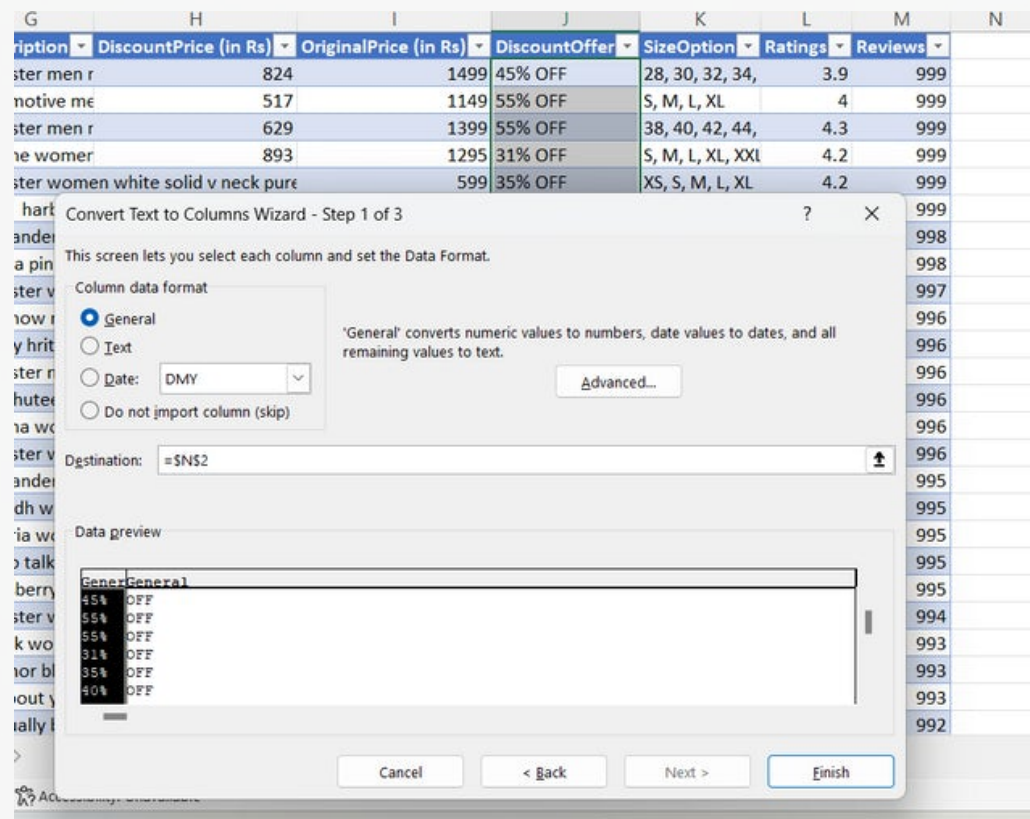
Microsoft Excel

No duplicate values found.

OK

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

We have to remove %OFF and HURRY from the discount offer for this use "TEXT TO COLUMN" and space delimited.



	N	O	P	Q	R	S	T
	Column	Column	Column				
	45% OFF		=IF(ISNUMBER([@Column2]),[@Column2],IF(
	55% OFF		ISBLANK([@Column1])," ",[@Column1]*				
	55% OFF		[@[OriginalPrice (in Rs)]]))				
	31% OFF		401.45				

Now the column will be standardized by converting values into numbers using the Formula displayed in the image

To Remove RS Use
CTRL + Key

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

Insert new column between "DISCOUNT OFFER" & "ORIGINAL PRIZE" then by using formula given in the image >we get "DISCOUNT OFFER"

	I	J
	Column4	OriginalPrice (in Rs)
824	=IF(ISBLANK([@[Discount Offer]),	
517	[@[OriginalPrice (in Rs)]]-[@[DISCOUNT	
629	AMOUNT]],@[Discount Offer])	
893		893
shirt	389.35	
	359.4	
599	599	
	585.9	
	#VALUE!	
le le	#VALUE!	
214	1214	

But there are some ERRORS in the new column to fix it use the formula displayed in the image

R	S	T
Column		
	=IF(ISERROR([@Column4])," ",	
	[@Column4])	
629		
893		
389.35		
359.4		
599		
585.9		

	S	T	U	V
al A	Column			
824	=IF([@[DISCOUNT AMOUNT]]=" ",			
517	AVERAGEIFS([DISCOUNT AMOUNT],			
629	[Category],[@Category]),@[DISCOUNT			
893	AMOUNT])			
9.35	209.65			
59.4	239.6			
599	899.4			
85.9	809.1			
	845.3686			
	1023.884			

After removing ERRORS there are some blank values to fill it use formula given in the image

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"

T	U	V	W
Column			
5	=SUBSTITUTE([@SizeOption], "",		
5	"NOT AVAILABLE")		
5			
5			
5			
6			
4			
1			
6 NULL VALUES			



R	S	T	U	V	W
Actual A	Column	Column	Column		
55				34, 36	
95					
45				44, 46, 48	
45				XXL	
55	38			XL	
.6	3			XL	
.4				36	
.1	5				
				XL	
				34, 36	
45				XXL	
				XL, XXL, 3XL, 4XL	
.3				XXL	
.5	12			XL	
.4	3			XL	
47				44	
41				XXL	
				XL, XXL, 3XL	
39	83			34	

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.

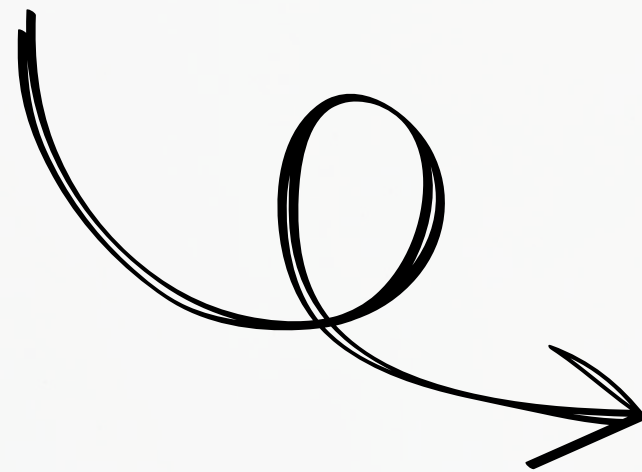
DATA ANALYSIS



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

To calculate overall average original price for RATINGS above 4 using the formula displayed in the image

average of original price rating > 4				
=ROUND(AVERAGEIF(Table1[Ratings], ">4", Table1[OriginalPrice (in Rs)]), 2)				



average of original price ratings>4

1675.49

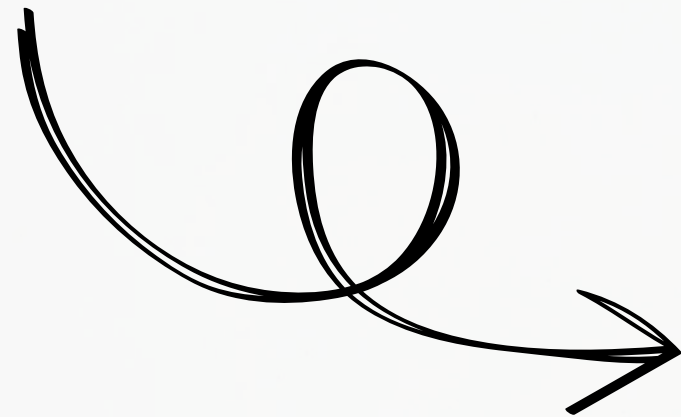


Count the number of products with a discount offer greater than 50% OFF

Use the COUNTIF Function to identify products with discount offer greater than 50%

products count where discount % is > 50%

```
=COUNTIF(Table1[Column1], ">50%")
```



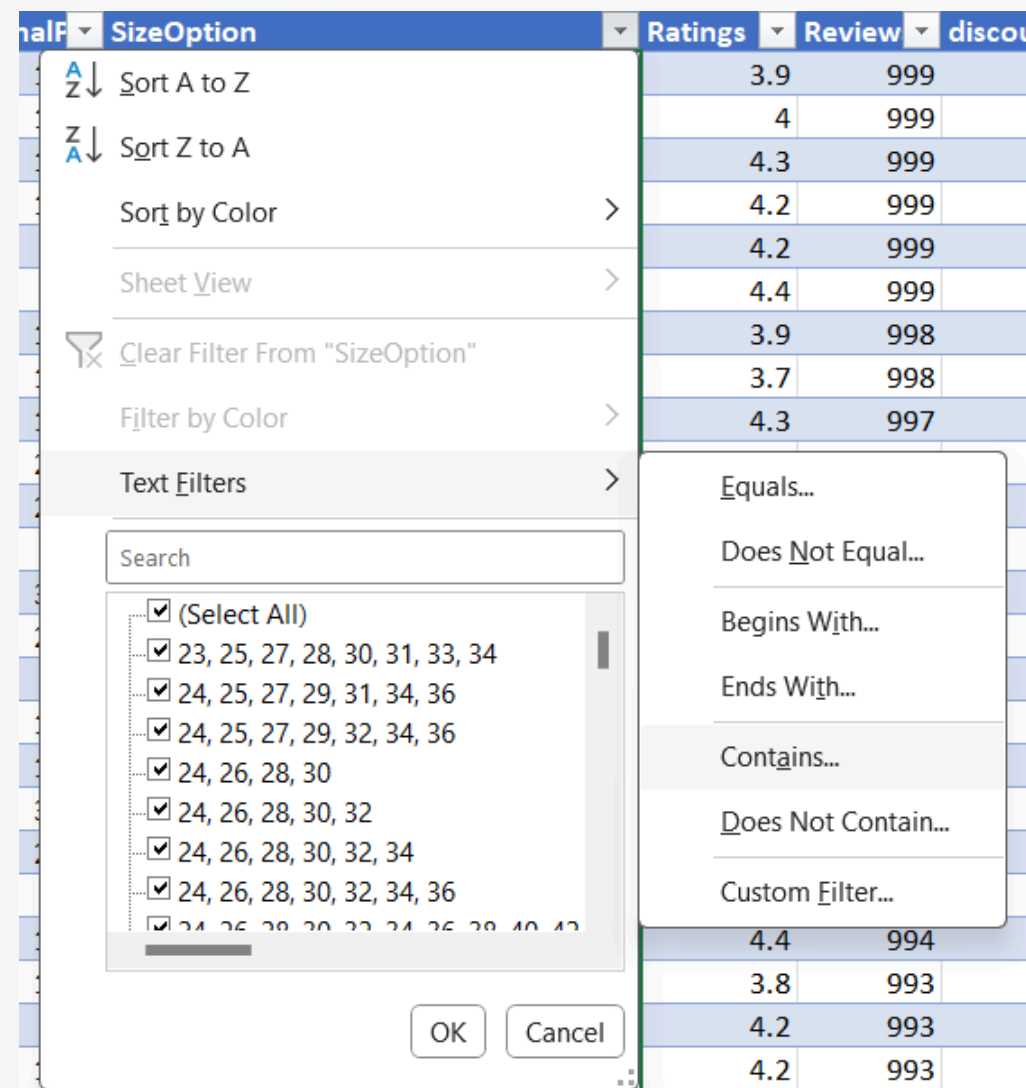
NO PRODUCT WERE DISCOUNT % IS > 50%

8194




Count the number of products available in size "M"

Here we are using the FILTER function of size option column to count the number of products with the condition where the size is "M"



Count the no of products available in size "M"

9835

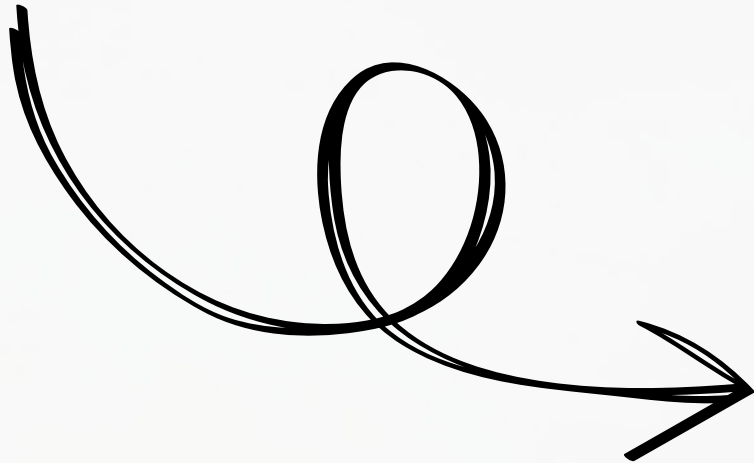


	SizeOption	Ratings	Review	discount_ame	Actual
49	S, M, L, XL	4	999	631.95	6
95	S, M, L, XL, XXL	4.2	999	401.45	4
99	XS, S, M, L, XL	4.2	999	209.65	2
99	XS, S, M, L, XL	4.4	999	239.6	
95	S, M, L, XL	3.7	998	809.1	
98	XS, S, M, L, XL	4.3	997		845.3
99	S, M, L, XL, XXL	4.4	996	1484.45	14
99	XS, S, M, L, XL, XXL, 3XL, 4XL	4.1	996		827.6
99	S, M, L, XL, XXL	4.2	996	2379.3	2
99	XS, S, M, L, XL	4.3	996	1249.5	1
99	XS, S, M, L, XL	4	996	479.4	
99	S, M, L, XL, XXL	4.2	995	1002.41	10
99	XS, S, M, L, XL, XXL, 3XL	4.3	995		1164
99	XS, S, M, L, XL	4.3	995	749.25	7
98	XS, S, M, L, XL	4.4	994		845.3
99	XS, S, M, L, XL, XXL	3.8	993	879.45	8
49	S, M, L, XL	4.2	993		535
99	XS, S, M, L, XL	4.2	993	749.5	
99	XS, S, M, L, XL, XXL	4.3	992	805.38	8
99	S, M, L, XL	4.2	991	740.43	7
99	XS, S, M, L, XL, XXL, 3XL	4	990	539.4	
98	L, M, S, XS, XL	4.2	990	799	
49	S, M, L, XL, XXL	4.4	990	1526.85	15
49	S, M, L, XL, XXL	4.3	990	1591.85	15
59	XS, S, M, L, XL	4.3	990	1627.45	16

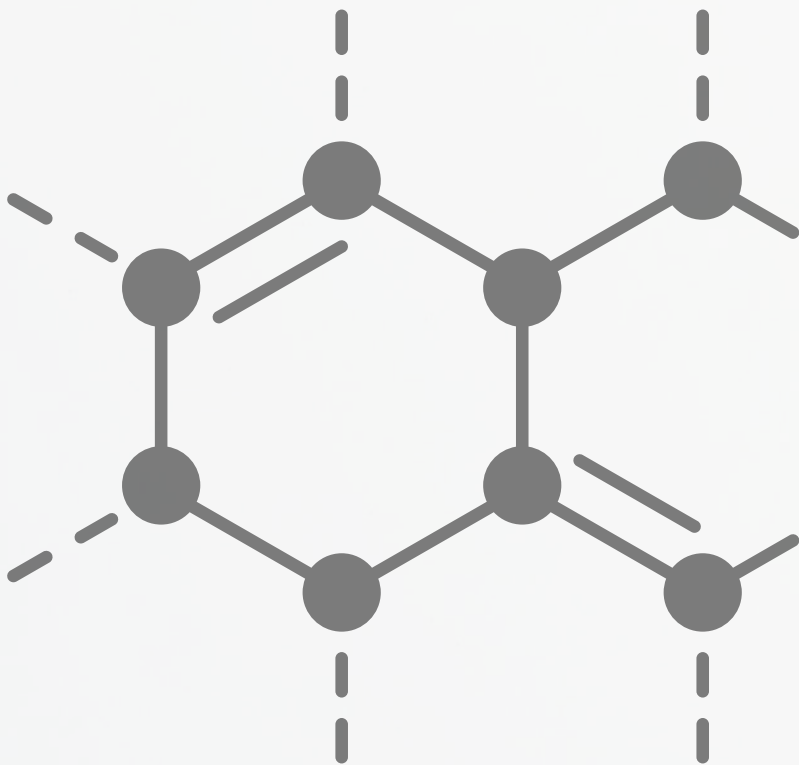
Count: 9835

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."

Create a new column by labeling the offer as "HIGH DISCOUNT" for greater than 50% & "LOW DISCOUNT" for less than 50% using the "IF" Function.



W	X	Y
Column35		
=IF([@[Discount Offer]]>50,"HIGH DISCOUNT","LOW DISCOUNT")		
HIGH DISCOUNT		
HIGH DISCOUNT		
LOW DISCOUNT		
LOW DISCOUNT		




DATA RETRIEVAL AND LOOKUP



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

We used the VLOOKUP Function to find the BRAND, PRIZE AND RATING of product ID:- 11226634



11226634		
BRAND	PRICE	RATING
Maniac	1199	3.9

BRAND

```
=VLOOKUP(S6,Table1[[Product_id]:[Column2]],{2},FALSE)
```

RATING

```
=VLOOKUP(S6,Table1[[Product_id]:[Column2]],{9},FALSE)
```

PRICE

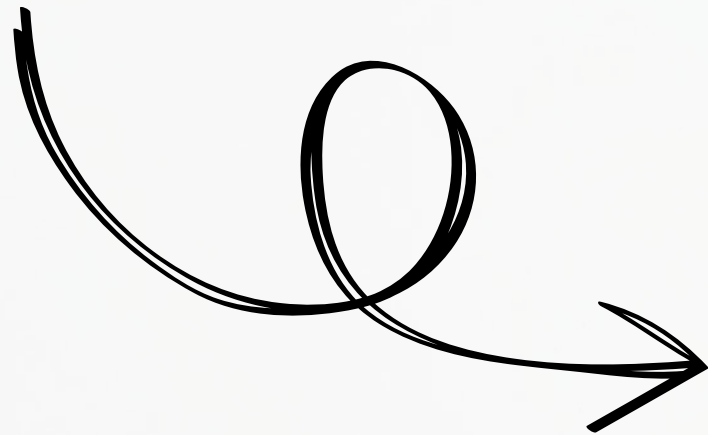
```
=VLOOKUP(S6,Table1[[Product_id]:[Column2]],{7},FALSE)
```

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

To find the discount prize for the product
ID:-6744434 using the index and MATCH Function

Index and Match for prouduct ID "6744434"

=INDEX(L:L,MATCH(Z1,B:B,0))

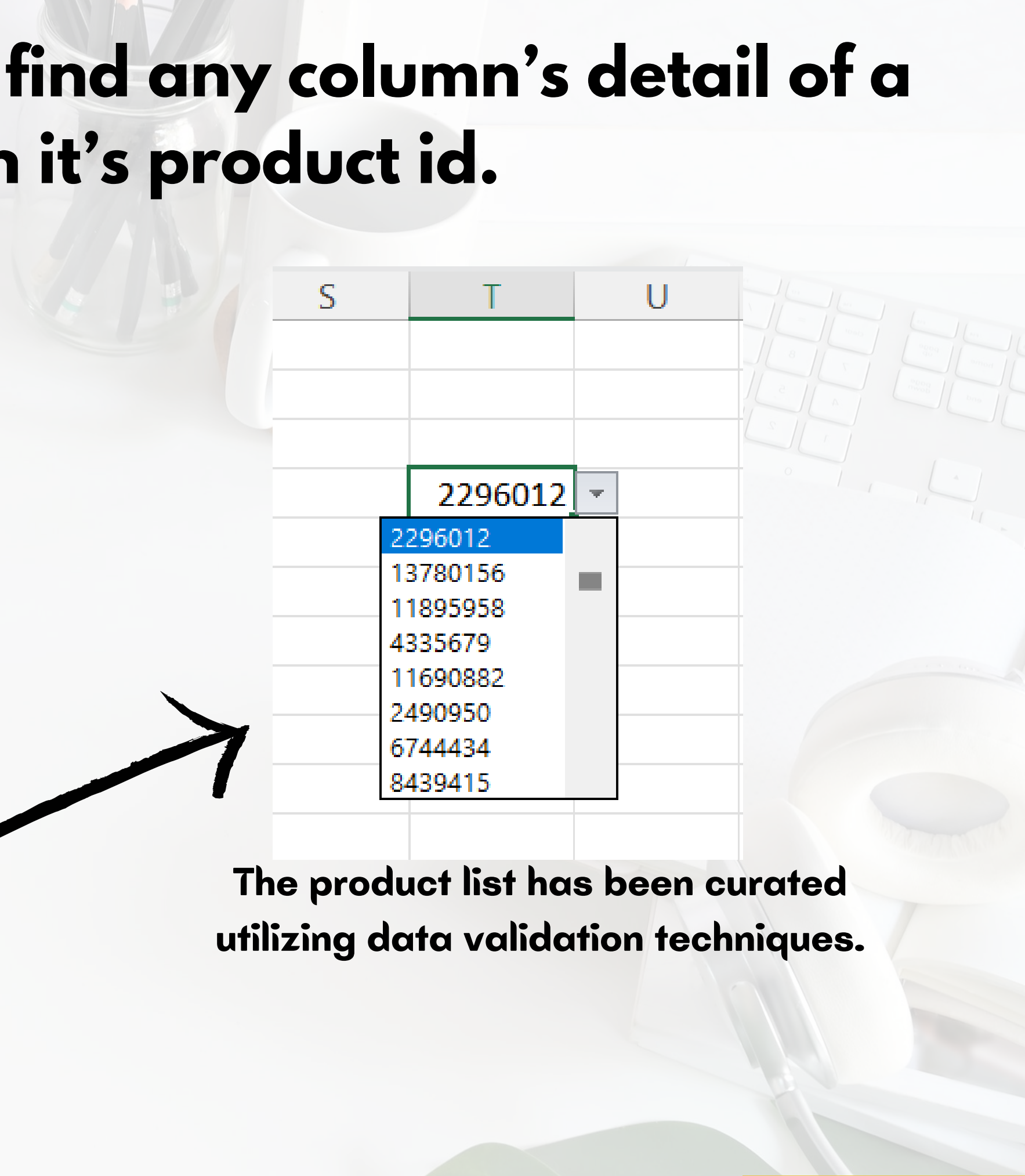


Index and Match for prouduct ID "6744434"

899.4



find any column's detail of a n it's product id.



S	T	U
	2296012	
	2296012	
	13780156	
	11895958	
	4335679	
	11690882	
	2490950	
	6744434	
	8439415	

The product list has been curated
utilizing data validation techniques.

DATA>DATA VALIDATION> SELECT LIST>FILL SOURCE>CLICK OK

[illegible]

The product list has been curated utilizing data validation techniques.

THANK YOU



www.linkedin.com/in/bhuvan-jari