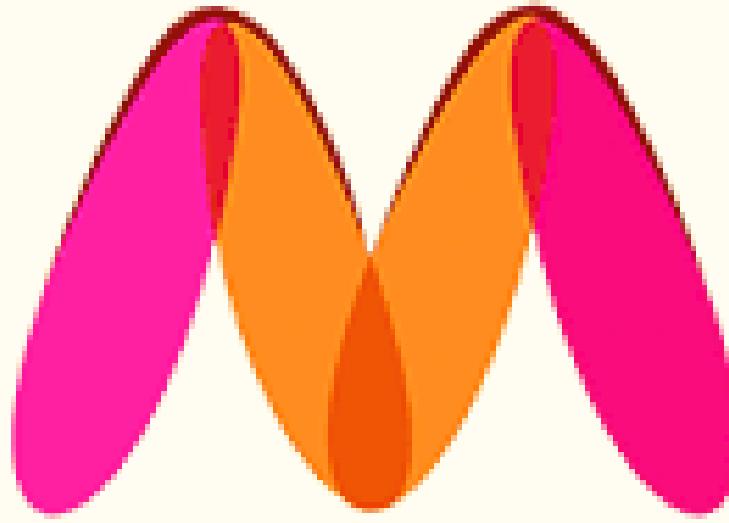




OSSYNTRA CASE STUDY



ANALYZE A DATASET OF VARIOUS APPAREL ITEMS TO GAIN INSIGHTS INTO PRICING, DISCOUNTS, RATINGS, AND AVAILABLE SIZES.



Data CLEANING & PREPARATION

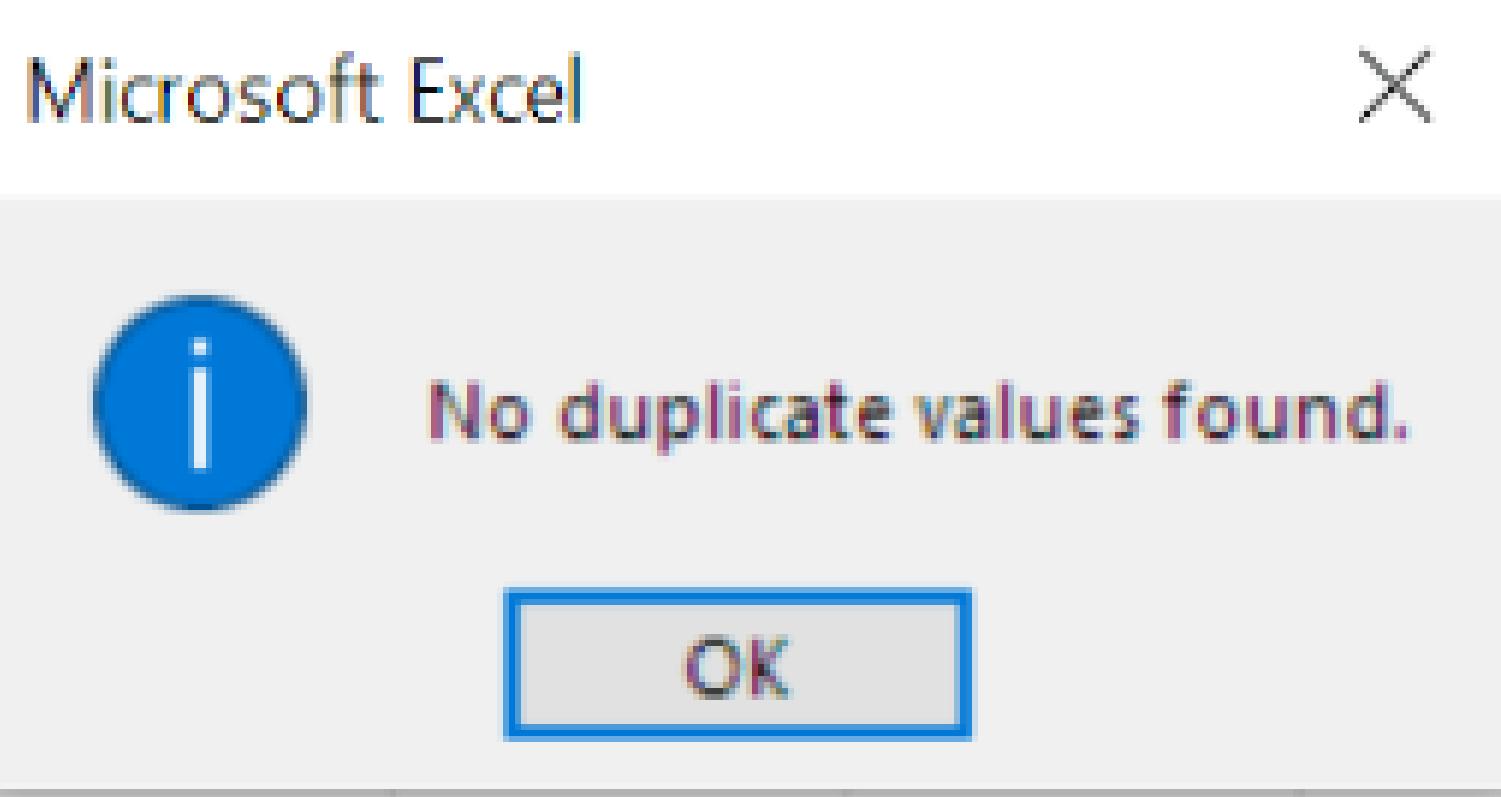


- 3** Check for duplicate values in your dataset and remove them.
- 4** Standardize the "DiscountOffer" column to a single format, ensuring all values are uniform.
- 5** Identify rows where both "DiscountPrice" and "DiscountOffer" are null and fill the "DiscountPrice" with the average discount price of the respective category.
- 6** Replace all null values in the "SizeOption" column with the text "Not Available."



ARE THERE ANY

duplicate values?



STANDARDIZATION



STANDARDIZED DISCOUNT COLUMN
674.55
631.95
769.45
401.45
389.35
209.65
899.4
809.1
0
0
1484.45
0
2379.3
1249.5
479.4

ALL VALUES IN NUMBERS ONLY.

EARLIER SOME WERE IN RS. FORMAT AND SOME WERE IN % FORMAT

1. REPLACE -> A. , HURRY*
B. OFF

2. # NEW COLUMN

- A. =SEARCH("RS.",J2)
- B. =ISNUMBER(SEARCH("RS.",J2))
- C. =IF(ISNUMBER(SEARCH("RS.",J2)),J2,I2*J2)
- D. COPY AND PASTE VALUES ONLY
- E. REPLACE -> RS.

IDENTIFYING ROWS WHERE BOTH "DISCOUNTPRICE" AND "DISCOUNTOFFER" ARE NULL AND FILL THE "DISCOUNTPRICE" WITH THE AVERAGE DISCOUNT PRICE OF THE RESPECTIVE CATEGORY.

Clipboard | Font | Alignment | Number | Styles

H6 : =IF(AND(ISBLANK(H2),ISBLANK(J2)),"MISSING","PRESENT")

	F	G	H	I
1	category_by_Gender	Description	DiscountPrice (in Rs)	OriginalPrice (in Rs)
2	Men	roadster men navy blue slim fit mid rise clean look jeans	824	
3	Men	locomotive men black white solid slim fit track pants	517	
4	Men	roadster men navy white black geometric printed sustainable casual shirt	629	
5	Women	zivame women black saree shapewear zi3023core0nude	893	
6	Women	roadster women white solid vneck pure cotton tshirt	PRESENT	
7	Women	mast harbour women yellow solid tank top	PRESENT	
8	Men	highlander men olive green slim fit solid regular trousers	599	
9	Women	mayra pink embroidered a line pure cotton top	PRESENT	

Font | Alignment | Number | Styles

=IF(AND(ISBLANK(H2),ISBLANK(J2)),AVERAGE([Category],D2,[DiscountPrice (in Rs)]),H2)

G	H	I
	DiscountPrice (in Rs)	OriginalPrice (in Rs)
mid rise clean look jeans	824	1499
slim fit track pants	517	1149
geometric printed sustainable casual shirt	629	1399
ewear zi3023core0nude	893	1295
ck pure cotton t shirt	824	599
id tank top	517	599
it solid regular trousers	629	1499

=IF(AND(ISBLANK(H2),ISBLANK(J2)),"MISSING","PRESENT")

=AVERAGEIF([Category],[@Category],[DiscountPrice (in Rs)])

G	H	I	J	K	L	M	N	O
	DiscountPrice (in Rs)	OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	STANDARDIZATION	AVERAGE
mid rise clean look jeans	824	1499	45%	28, 30, 32, 34,	3.9	999	674.55	878
slim fit track pants	517	1149	55%	S, M, L, XL	4	999	631.95	878
geometric printed sustainable casual shirt	629	1399	55%	38, 40, 42, 44,	4.3	999	769.45	667
ewear zi3023core0nude	893	1295	31%	S, M, L, XL, XXL	4.2	999	401.45	624
ck pure cotton t shirt	PRESENT	599	35%	XS, S, M, L, XL	4.2	999	209.65	#DIV/0!
id tank top	PRESENT	599	40%	XS, S, M, L, XL	4.4	999	239.6	#DIV/0!
it solid regular trousers	599	1499	60%	30, 32, 34, 36	3.9	998	899.4	878
ure cotton top	PRESENT	1395	58%	S, M, L, XL	3.7	998	809.1	#DIV/0!

=IF(AND(ISBLANK(H6),ISBLANK(J6)),AVERAGE([CATEGORY],D6,[DISCOUNTPRICE (IN RS)]),H6)

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."

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

1609.409217

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

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

781

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

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

1007

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

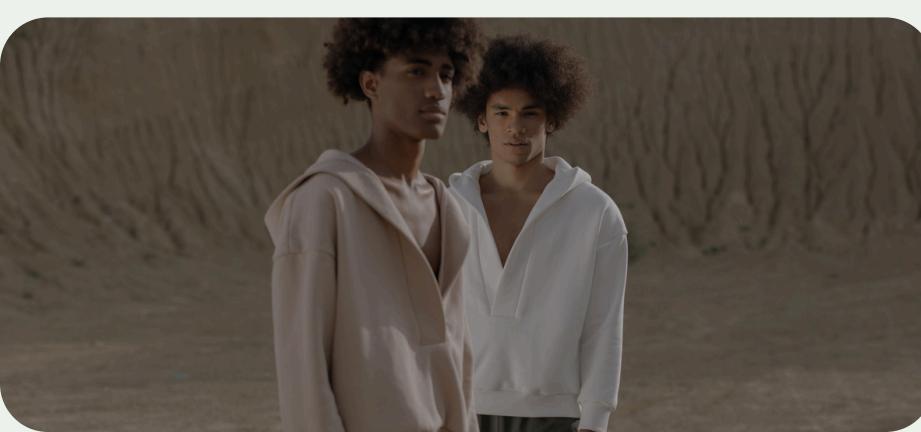
4. Create a new column to label the products as "High Discount" if the discount offer is greater than 50% OFF, otherwise label them as "Low Discount."

COL. O

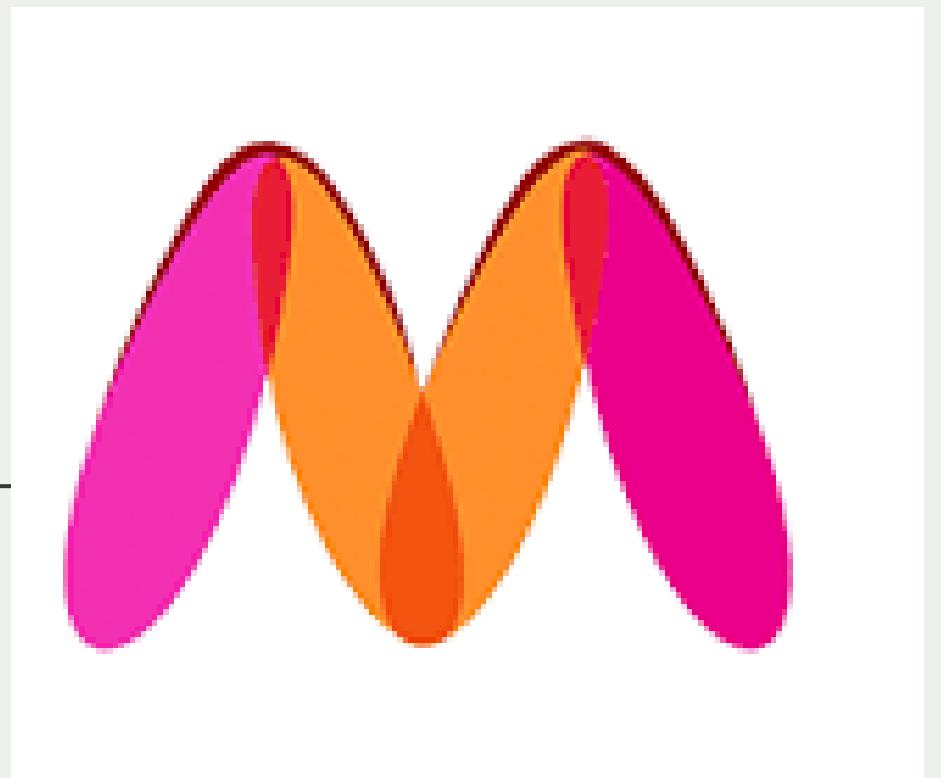


PAUCEK AN

DISCOUNT LABEL
LOW DISCOUNT
HIGH DISCOUNT
HIGH DISCOUNT
LOW DISCOUNT
HIGH DISCOUNT



DATA **RETRIEVAL AND LOOKUP**

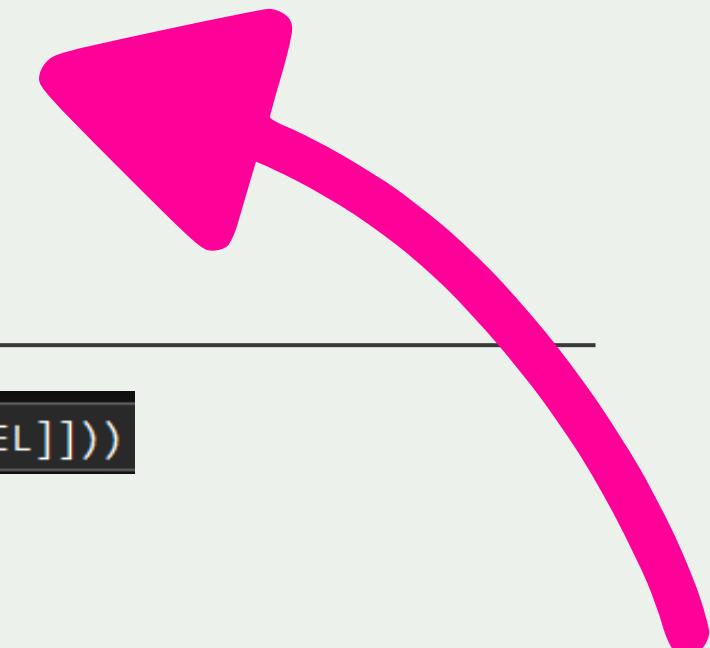


1. USE VLOOKUP/XLOOKUP TO FIND THE PRODUCT BRAND, PRICE, AND RATING OF THE PRODUCT WITH PRODUCT_ID "11226634".
2. FIND THE "DISCOUNTPRICE" FOR THE PRODUCT WITH THE PRODUCT ID "6744434" USING THE INDEX AND MATCH FUNCTIONS.
3. UTILIZE NESTED XLOOKUP TO FIND ANY COLUMN'S DETAIL OF A PRODUCT WITH IT'S PRODUCT ID.

5. Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product_id "7372858".

	7372858
Ratings	
	4.1

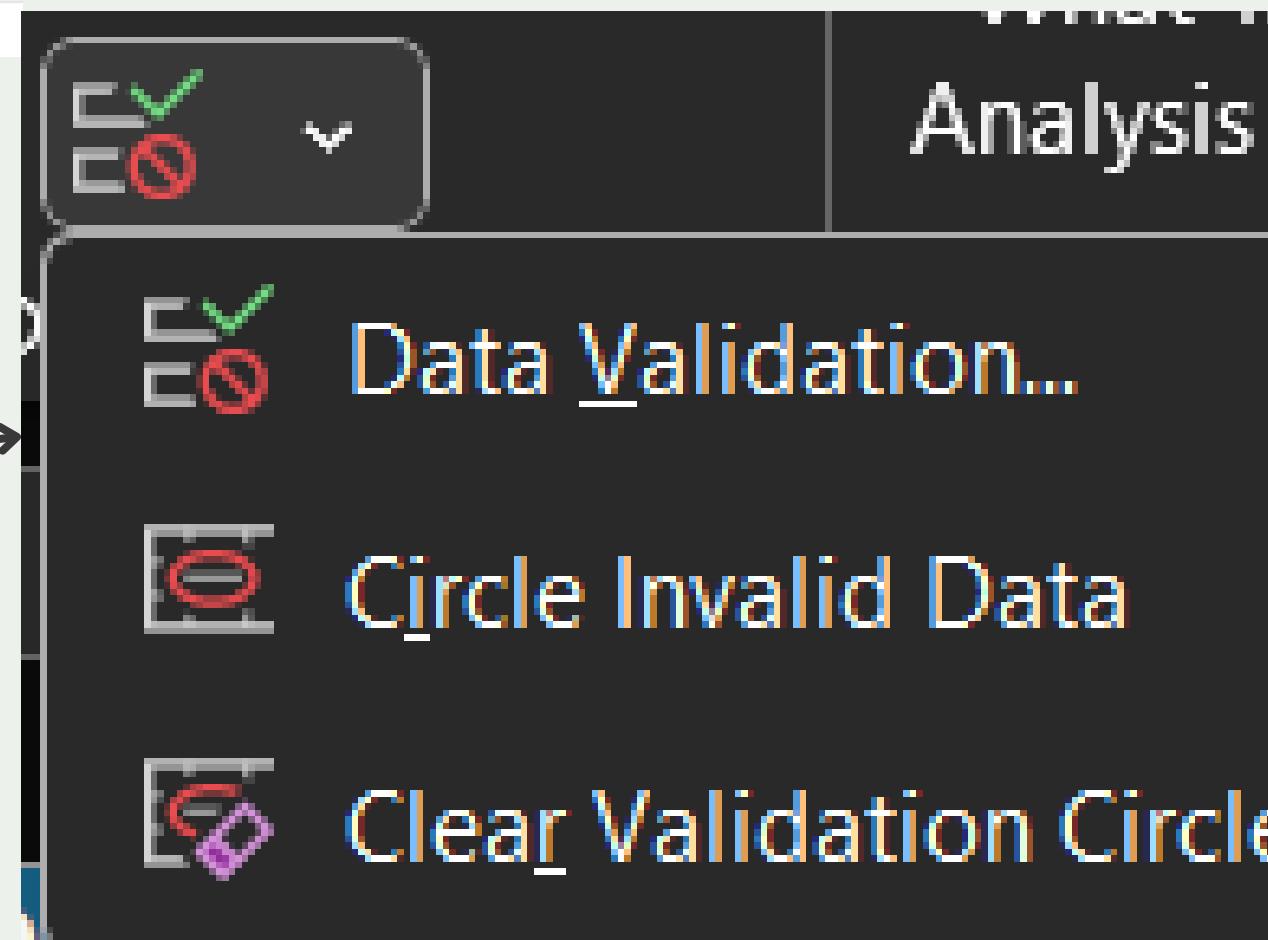
```
=XLOOKUP(G1513,Table1[Product_id],XLOOKUP(G1515,Table1[#Headers],[BrandName]:[DISCOUNT_LABEL]),Table1[[BrandName]:[DISCOUNT_LABEL]]))
```



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

			6744434
index & match		index	match
	629	629	7

```
=INDEX(Table1[DiscountPrice (in Rs)],MATCH(I1521,Table1[Product_id],0))
```



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

	10145883
Individual_category	
	shirts

```
=XLOOKUP(G1529,Table1[Product_id],XLOOKUP(G1531,Table1[#Headers],Table1[[URL]:[DISCOUNT_LABEL]]))
```

10145883

shirts



Thank
you.