



# ANALYSIS OF MYNTRA PROJECT

**Present By : Disha Sindhii**

# INTRODUCTION

Myntra is an Indian e-commerce company that specializes in and lifestyle products. Founded in 2007, it offers a wide range of clothing, accessories, and beauty products from various brands. Myntra is Known for its user-friendly interface, discount and unique shopping experiences, including curated collections and personalized recommendations. it's also Popular for its mobile. app, making online shopping convenient for users.

# PROBLEMS

A. Data Cleaning and Preparation

B. Data Analysis

C. Data Retrieval and Lookup



# A. Data Cleaning and Preparation



Q 1. CHECK FOR DUPLICATE VALUES IN YOUR DATASET AND REMOVE

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

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

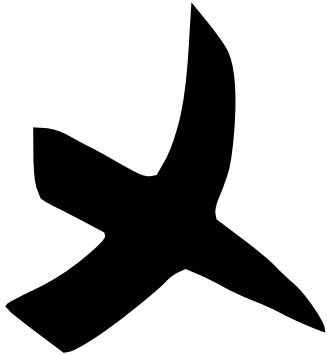
☒ Category

☒ Individual\_category

OK

Cancel

|                   |     |         |              |              |     |     |
|-------------------|-----|---------|--------------|--------------|-----|-----|
| roadster          | 834 | 1499    | 45% OFF      | 39, 40, 42,  | 4.2 | 999 |
| locomotive        |     |         |              |              |     | 999 |
| roadster          |     |         |              |              |     | 999 |
| zivame w          |     |         |              |              |     | 999 |
| roadster          |     |         |              |              |     | 999 |
| mast har          |     |         |              |              |     | 999 |
| highlander        |     |         |              |              |     | 998 |
| mayra pi          |     |         |              |              |     | 998 |
| roadster          |     |         |              |              |     | 997 |
| herenow           |     |         |              |              |     | 996 |
| hrx by hr         |     |         |              |              |     | 996 |
| roadster          |     |         |              |              |     | 996 |
| anubhute          |     |         |              |              |     | 996 |
| athena w          |     |         |              |              |     | 996 |
| roadster women ma | 799 | 60% OFF | XS, S, M, L, | 4            |     | 996 |
| highlander        | 516 | 1099    | 53% OFF      | 39, 40, 42,  | 4.2 | 995 |
| vishudh w         | 696 | 1699    | 59% OFF      | S, M, L, XL, | 4.2 | 995 |



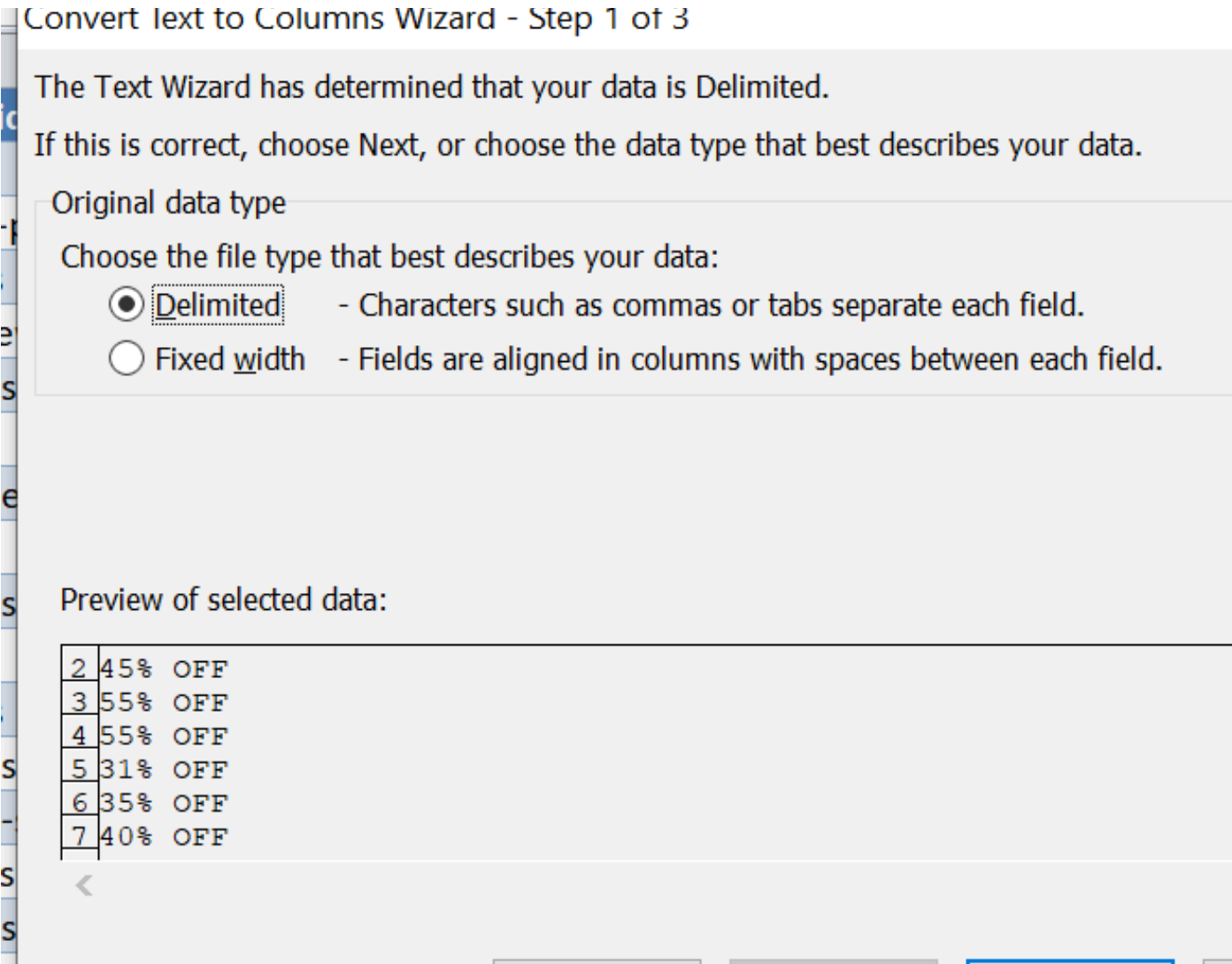
|         |      |         |              |     |
|---------|------|---------|--------------|-----|
| 893     | 1295 | 31% OFF | S, M, L, XL, | 4.2 |
| en whi  | 599  | 35% OFF | XS, S, M, L, | 4.2 |
| wome    | 599  | 40% OFF | XS, S, M, L, | 4.4 |
| 599     |      |         |              | 3.9 |
| nbroide |      |         |              | 3.7 |
| en pac  |      |         |              | 4.3 |
| blue s  |      |         |              | 3.5 |
| 1214    |      |         |              | 4.4 |
| navy b  | 699  |         | XS, S, M, L, | 4.1 |
| 1019    | 3399 | 70% OFF | S, M, L, XL, | 4.2 |
| n black | 2499 | 50% OFF | XS, S, M, L, | 4.3 |
| en ma   | 799  | 60% OFF | XS, S, M, L, | 4   |
| 516     | 1099 | 53% OFF | 39, 40, 42,  | 4.2 |
| 696     | 1699 | 59% OFF | S, M, L, XL, | 4.2 |
| n green | 3999 |         | XS, S, M, L, | 4.3 |

Microsoft Excel

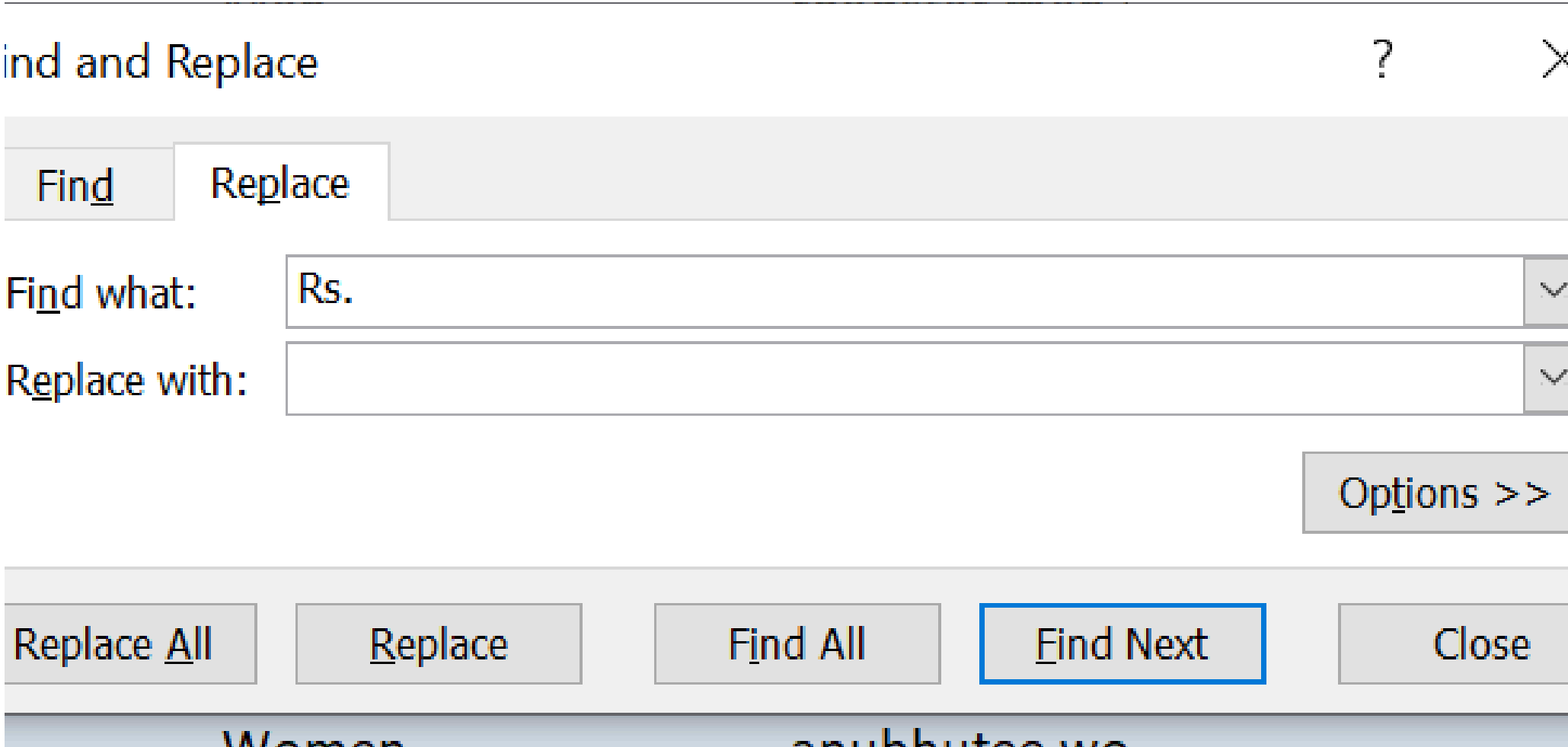
No duplicate values found.

OK

Q 2. Standardize the “Discountoffer” column to a single format, ensuring all values are uniform.



Select the “Discountoffer”  
Go to Data->Text to Column  
and replace all rupees (Rs.)  
Select the “Discountoffer”  
and press ctrl + F



## Standardize “DiscountOffer” column to uniform format.

New, the Column will be standardized by converting percentage values into number using the formula displayed in the image

| Offer3 | Discour   |
|--------|---|
|        | =IF(ISNUMBER(<br>[@DiscountOffer3]),<br>[@DiscountOffer3],<br>[@DiscountOffer3],<br>[@DiscountOffer2]*<br>[@[OriginalPrice (in Rs)<br>]]) |
|        | gical_test, [value_if_true], [value_if_false])  |
|        | 809.1   |
|        | 0   |
|        | 0   |
|        | 1484.45   |
|        | 0   |
|        | 2379.3  |

| Discour  |
|--|
| =IF(ISNUMBER([@DiscountOffer3]),[@DiscountOffer3],IF(ISBLANK([@DiscountOffer2]),<br>" ",[@DiscountOffer2]*[@[OriginalPrice (in Rs)]])) |
| IF(logical_test, [value_if_true], [value_if_false])  |
| 401.45   |
| 209.65   |
| 239.6  |
| 899.4  |
| 809.1  |

Convert all “0” Values to  
Blank Values



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

| Column1   | OriginalPrice | SizeOpt      | Ratings | Review | Discount |
|---|---------------|--------------|---------|--------|----------|
| =IF(ISBLANK([@[Actual Price after discount]]),[@[OriginalPrice (in Rs)]]-[@[Discount_amount]],[@[Actual Price after discount]]) |               |              |         |        |          |
| IF(logical_test, [value_if_true], [value_if_false])   | 4             | 999          | 631.95  |        |          |
| 629   | 1399          | 38, 40, 42,  | 4.3     | 999    | 769.45   |
| 893   | 1295          | S, M, L, XL, | 4.2     | 999    | 401.45   |
| 389.35  | 599           | XS, S, M, L, | 4.2     | 999    | 209.65   |
| 359.4   | 599           | XS, S, M, L, | 4.4     | 999    | 239.6    |
| 599   | 1499          | 30, 32, 34,  | 3.9     | 998    | 899.4    |
| 585.9   | 1395          | S, M, L, XL  | 3.7     | 998    | 809.1    |

Now Finding the "Actual Price"

| N        | O                    | P | Q |
|----------|----------------------|---|---|
| Discount | Column1              |   |   |
| 674.55   | =ISERROR([@Column1]) |   |   |
| 631.95   | ISERROR(value)       |   |   |
| 769.45   | FALSE                |   |   |
| 401.45   | FALSE                |   |   |
| 209.65   | FALSE                |   |   |
| 239.6    | FALSE                |   |   |
| 899.4    | FALSE                |   |   |

| Discount | Column1   |  |  |
|----------|---|--|--|
| 74.55    | =IF(ISERROR([@Column1])," ",[@Column1])                 |  |  |
| 31.9     | Dis IF(logical_test, [value_if_true], [value_if_false]) |  |  |
| 59.4     | (Showing All)   |  |  |
| 01.45    | 893   |  |  |
| 09.65    | 389.35  |  |  |
| 239.6    | 259.4   |  |  |

Remove all Error's



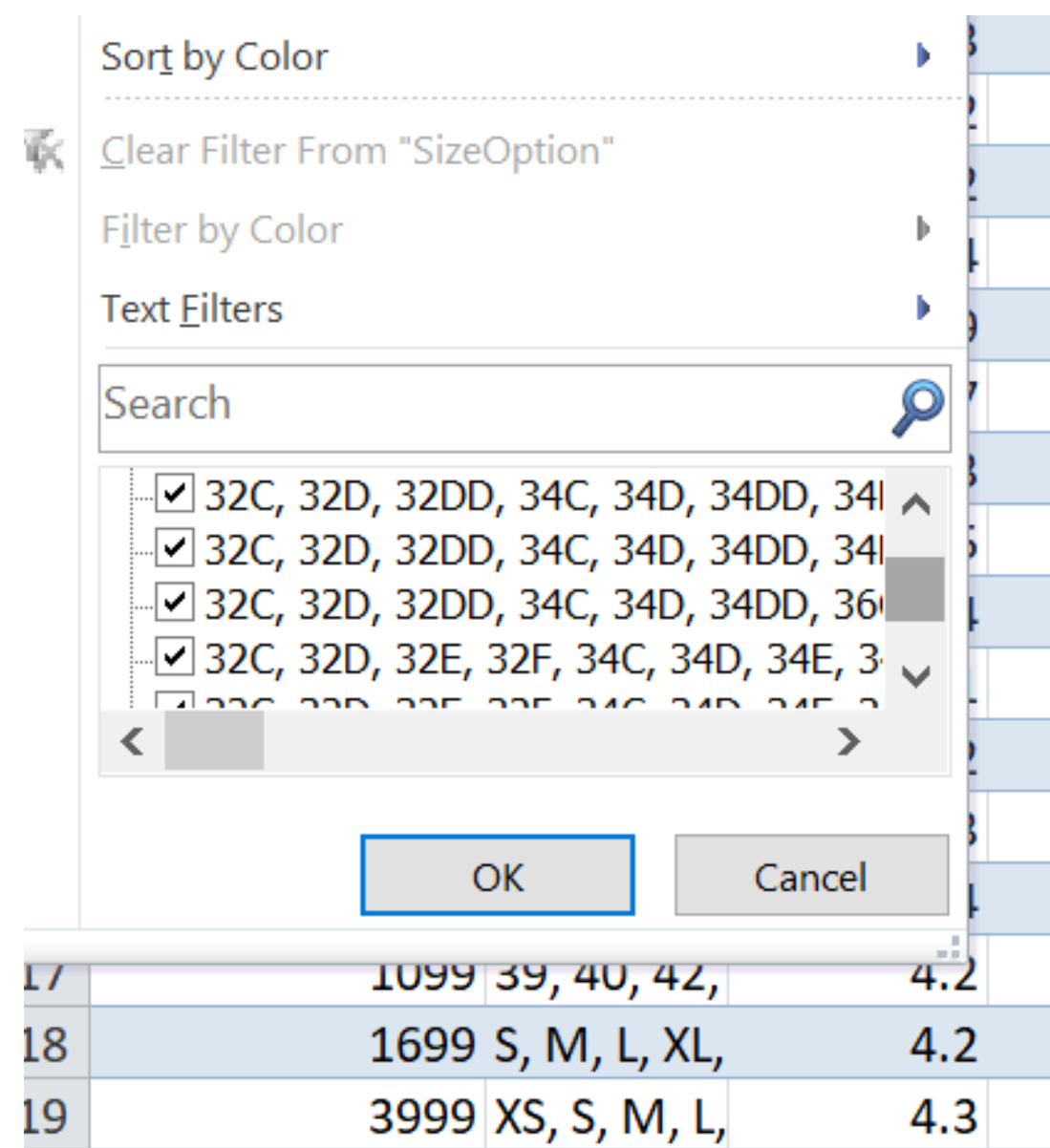
**The “Discount price” with the average discount price of the respective category**

|      | Columnr   |
|------|---|
| 824  | =IF([@[Discount_amount]]=" ",AVERAGEIFS([Discount_amount],[Category],[@Category]),[@[Discount_amount]]) |
| 517  | 631.95  |
| 629  | 769.45  |
| 893  | 401.45  |
| 9.35 | 209.65  |

|     |  |  |  |  |  |  |  |  |  |  |
|-----|--|--|--|--|--|--|--|--|--|--|
| 45% | =IF(Table1[@[Discount Amount]]=" ",Table1[@[OriginalPrice (in Rs)]]-Table1[@[DiscountPrice (in Rs)3]], |  |  |  |  |  |  |  |  |  |
| 55% | Table1[@[Discount Amount]])  |  |  |  |  |  |  |  |  |  |
| 55% | 769.45   |  |  |  |  |  |  |  |  |  |
| 31% | 401.45   |  |  |  |  |  |  |  |  |  |
| 35% | 209.65   |  |  |  |  |  |  |  |  |  |
| 40% | 239.6  |  |  |  |  |  |  |  |  |  |
| 60% | 899.4  |  |  |  |  |  |  |  |  |  |
| 58% | 809.1  |  |  |  |  |  |  |  |  |  |
| 88% | 127.8412   |  |  |  |  |  |  |  |  |  |
| 38% | 1699.797   |  |  |  |  |  |  |  |  |  |
| 55% | 1484.45  |  |  |  |  |  |  |  |  |  |
| 32% | -226.725   |  |  |  |  |  |  |  |  |  |
| 78% | 2279.2   |  |  |  |  |  |  |  |  |  |

## Now Finding the “Discount Amount”

Q 4. Replace all null values in the “SizeOption” column with the text “Not Available



**“SizeOption”  
Column has no null  
values**



# DATA ANALYSIS

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

|  |  |  |  |  |  |
|--|--|--|--|--|--|
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  | Average of Original Price whose rating is more than 4                      |  |  |  |  |
|  | =ROUND(AVERAGEIF(Table1[Ratings], ">4", Table1[OriginalPrice (in Rs)]), 2) |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**Calculate the Average original price for rating above 4 using the formula**

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

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

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

COUNTIF(**range**, criteria)

**We Will Utilize the Countifs Function to identify products with discount offers exceeding 50%**

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

Sort A to Z

Sort Z to A

Sort by Color

Clear Filter From "SizeOption"

Filter by Color

Text Filters

Search

☒ (Select All)

☒ 22

☒ 25

☒ 26

☒ 28

OK

Cancel

Equals...

Does Not Equal...

Begins With...

Ends With...

Contains...

Does Not Contain...

Custom Filter...

Count: 308460

Sum: 0

100%

1

36°C

^

📹

🔄

📶

🔌

🔊

EN

IT

Custom AutoFilter

Show rows where:

SizeOption

contains

M

☒ And

☐ Or

Use ? to represent any single character

Use \* to represent any series of characters

OK

Cancel

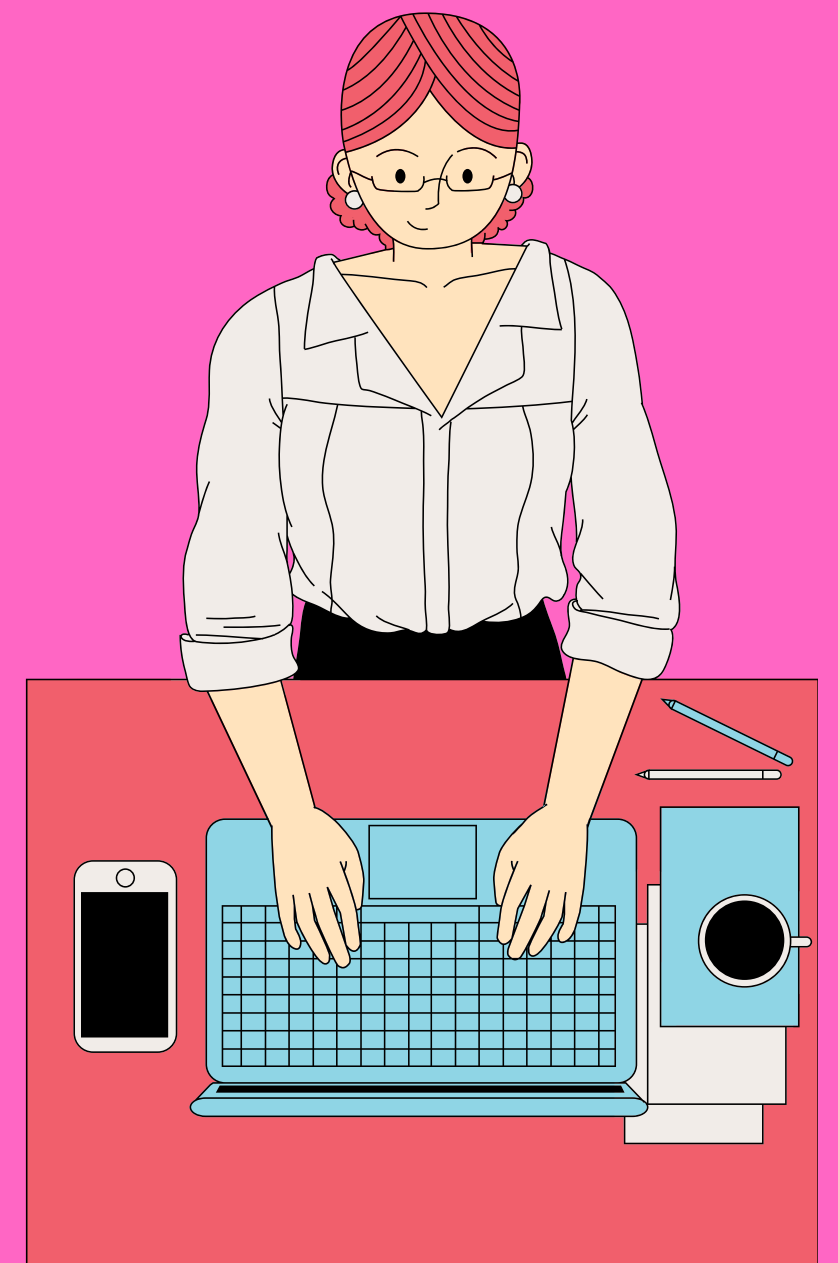
**Q 4. Create a new column to label the products as "High Discount" if the discount offer is greater than 50% OFF, otherwise label them as "Low Disc"**

| Column1 |  |  |  |  |
|---------|--|--|--|--|
| %       | =IF([@[ discount offer]]>0.5,"High discount","Low discount") |  |  |  |
| %       | IF(logical_test, [value_if_true], [value_if_false])          |  |  |  |

**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



**Q 1. Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product\_id "11226634"**

[illegible]

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

Index and Match for Product Id "6744434"

=INDEX(Table1[DiscountPrice (in Rs)3],MATCH(P7,Table1[Product\_id],0))

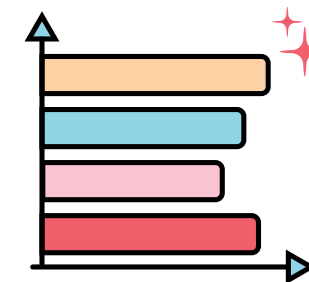
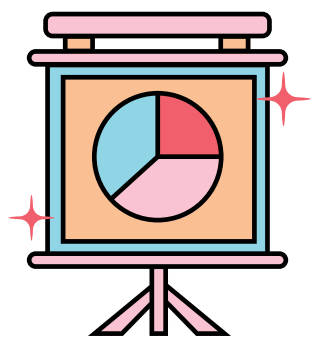
INDEX(array, row\_num, [column\_num])

INDEX(reference, row\_num, [column\_num], [area\_num])

Index and Match for Product Id "6744434"

6744434

899.4



**THANK  
YOU**

