

Fueling the Future, Powering Progress

# MYNTRA FASHION CLOTHING

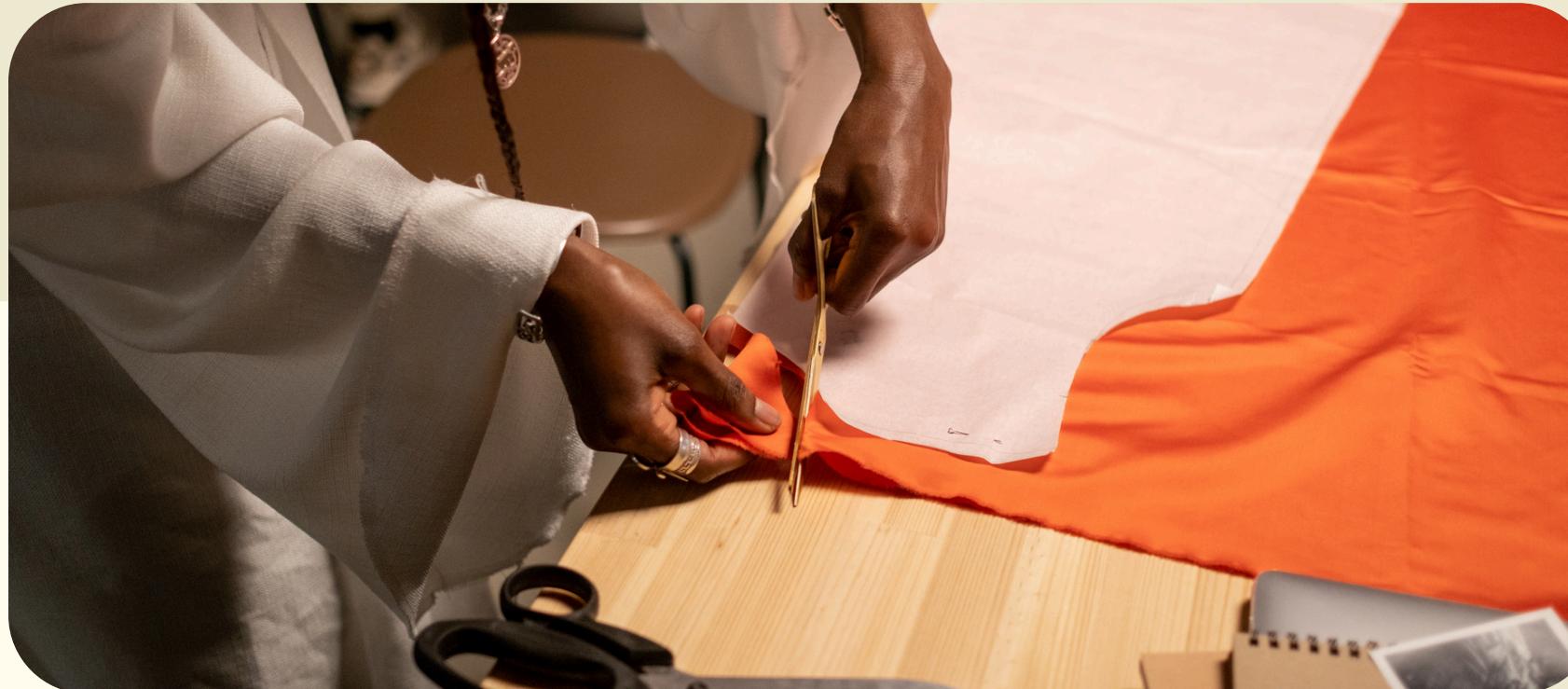
[www.myntra.com](http://www.myntra.com)



# About Us

---

This project is based on a dataset from Myntra, a leading online fashion retailer in India. The aim of the project is to analyze apparel data to derive meaningful insights into pricing, discounts, ratings, and availability of products. The project focuses on data cleaning, preparation, and analysis using Excel functions.



## ***Objectives of the Project:***

- Clean and prepare the dataset by removing duplicates and handling missing values.
- Standardize data formats such as discount offers and pricing.
- Perform analysis to understand product pricing trends, discount patterns, and availability.
- Use lookup functions to retrieve product-specific details efficiently.
- Label and categorize products based on discount levels for better decision-making.



# MYNTRA Company Vision and Mission

---

## Vision

Mynta envisions a world that is "more stylish, colorful, and happier." The company's strategy revolves around this vision—each department and initiative is aligned to make fashion an uplifting and joyful experience for everyone.

## Mission

Mynta's mission is to "leverage technology to democratize fashion and lifestyle—to help people look and feel good." This reflects its commitment to merging technology with fashion accessibility, ensuring consumers across the country can enjoy a seamless and personalized shopping experience

# Our Products

---

- 1. Fashion Apparel

**Men, Women & Kids:**

Ethnic Wear: Sarees, kurtas, lehengas

Formal & Casual: T-shirts, shirts, trousers, skirts, dresses, blazers

Activewear & Outerwear: Gym wear, sports apparel, jackets, coats, sweaters

Innerwear & Sleepwear: Lingerie, pajamas, loungewear, thermals

- 2. Footwear

Casual: Sneakers, sandals, loafers

Formal: Oxfords, brogues, heels, boots

Ethnic: Mojaris, juttis

Sports: Running and training shoes

- 3. Accessories & Jewelry

Accessories: Bags, wallets, belts, caps, scarves, stoles, sunglasses, ties, socks

Jewelry: Necklaces, bracelets, earrings, rings, anklets

Watches: Smartwatches, analog/digital watches



## 5. Sport & Active Essentials

Activewear, gym apparel, yoga pants, sports bras  
Sports accessories like water bottles and gym bags

## 4. Beauty & Personal Care

Skincare: Serums, moisturizers, sunscreens

Makeup: Lipsticks, foundations, eyeliners, palettes

Haircare and Grooming: Shampoos, oils, shaving kits

Fragrances: Perfumes, body mists, deodorants



# OBJECTIVES OF THE PROJECT:

Clean and prepare the dataset by removing duplicates and handling missing values.

Standardize data formats such as discount offers and pricing.

Perform analysis to understand product pricing trends, discount patterns, and availability.

Use lookup functions to retrieve product-specific details efficiently.

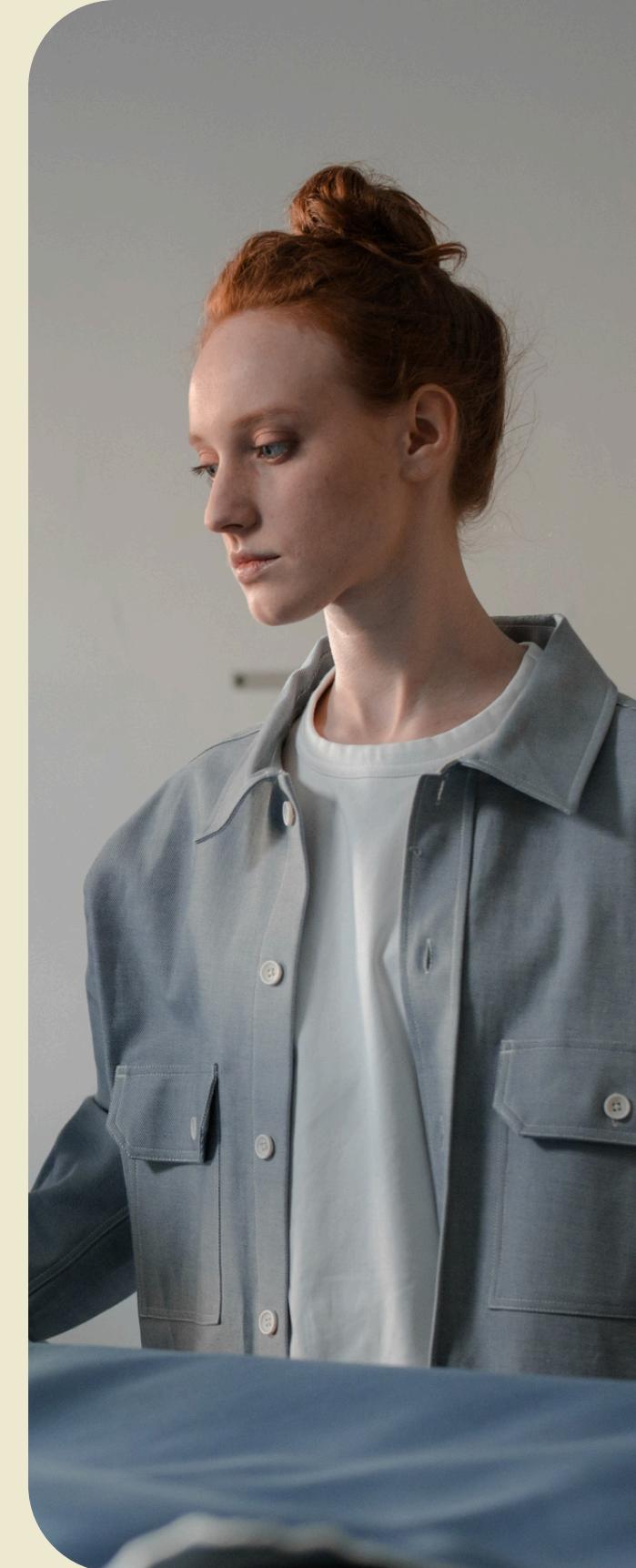
Label and categorize products based on discount levels for better decision-making.



# PROJECT TASKS AND QUESTIONS

Remove duplicates and standardize discount columns.

1. Replace missing discount values with category-wise averages.
2. Handle missing size information.
3. Calculate the average price of products with high ratings (>4).
4. Identify products with discounts above 50% and categorize them as High Discount or Low Discount.
5. Count available products by size (e.g., size "M").
6. Use VLOOKUP, XLOOKUP, INDEX & MATCH to find product details such as brand, price, rating, and discount.



## Project Questions

### A. 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."

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

Duplicates based on unique identifiers : In this data set, the product id serves as a unique identifier

- a. select the unique column (product id)
- b. Go to “ Home” tab and click on “Conditional Formatting”
- c. select “Highlight cells Rules” , then select “ Duplicates values”
- d. Go to “Data” tab and select “ Remove Duplicates”

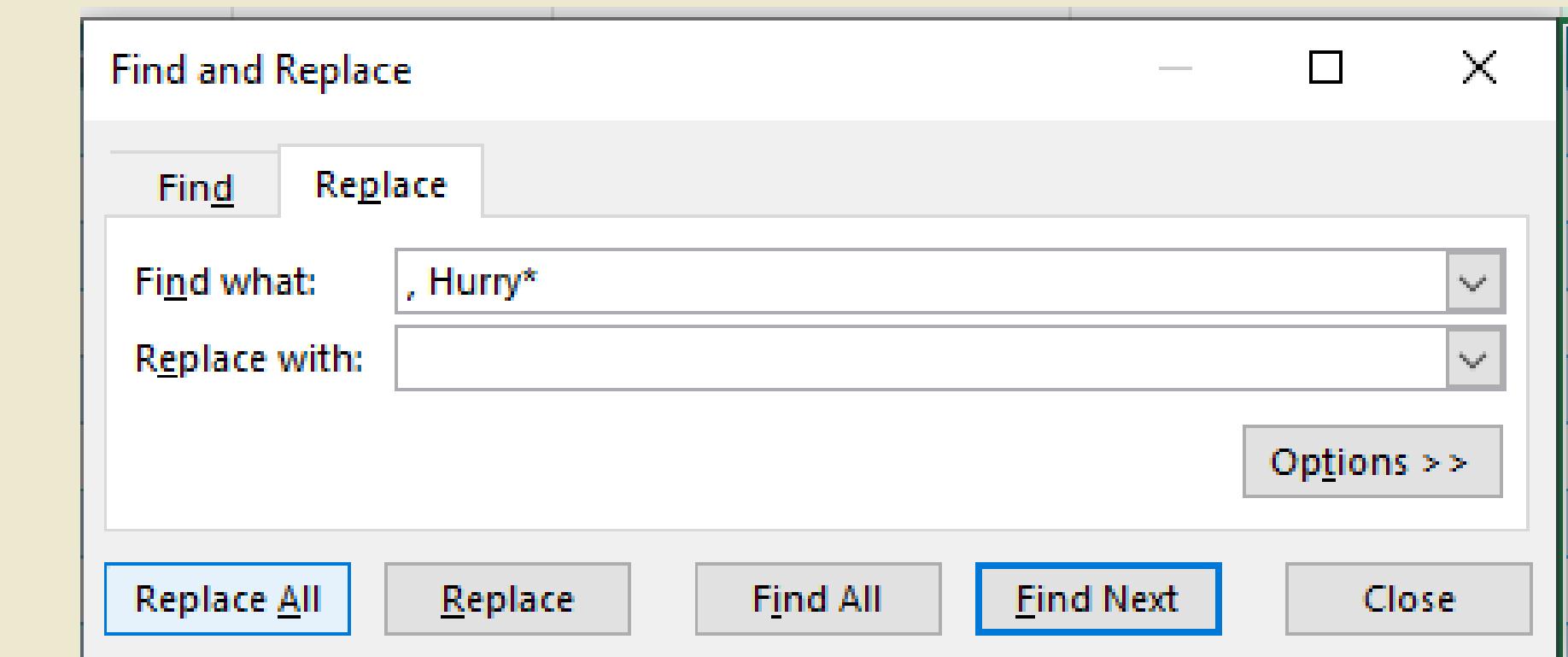
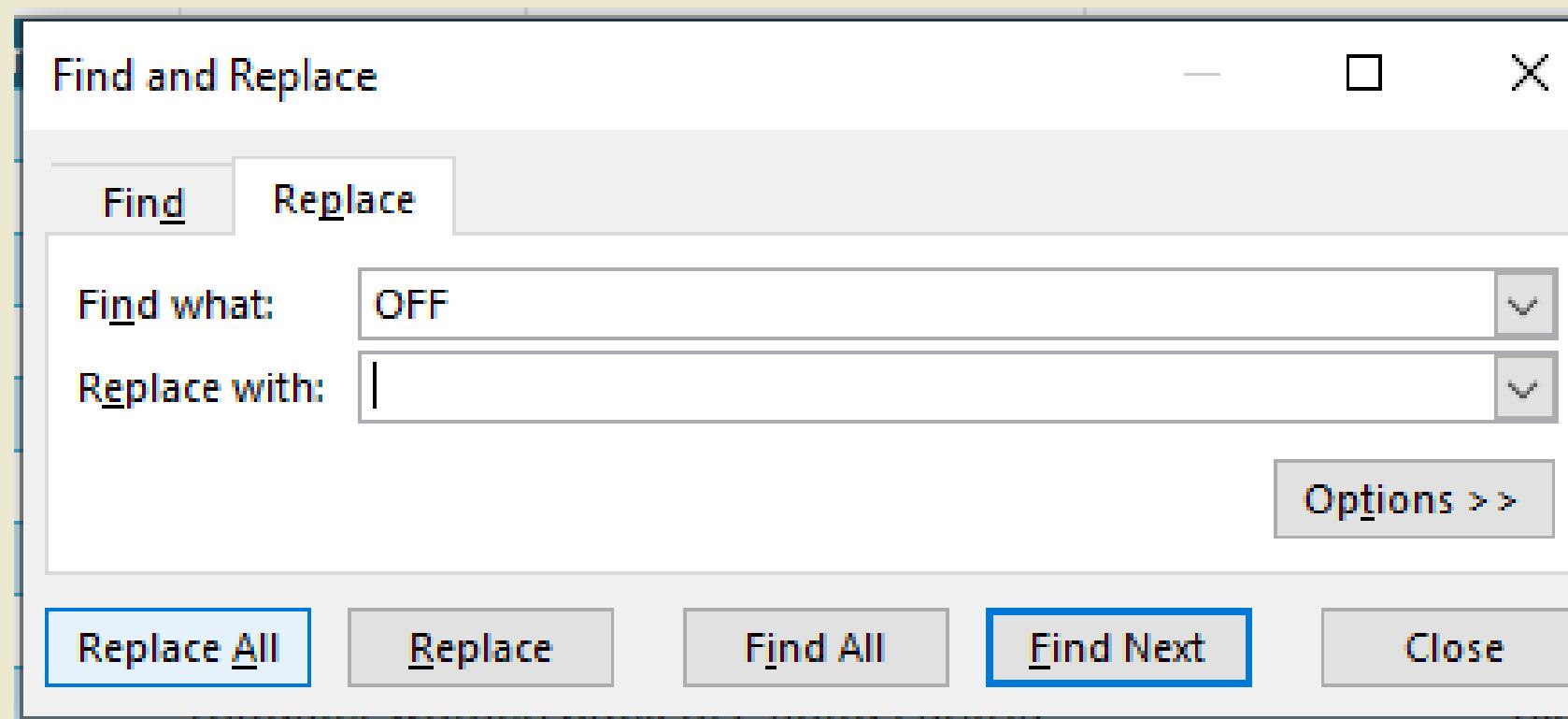
Conclusion : There are no duplicates in this data set

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

Step 1. Use Ctrl+H to open “ Find and Replace” . Replace  
“ Off”, “ Hurry” with desired text.

Step 2. Use the Formula:

=IF(ISNUMBER(SEARCH("Rs",[@Discounter])),  
[@Discounter],[@Discounter]\*[@[OriginalPrice(in Rs)]])



=IF(ISNUMBER(SEARCH("Rs",[@Discounter])),[@Discounter],[@Discounter]\*[@[OriginalPrice (in Rs)]])

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

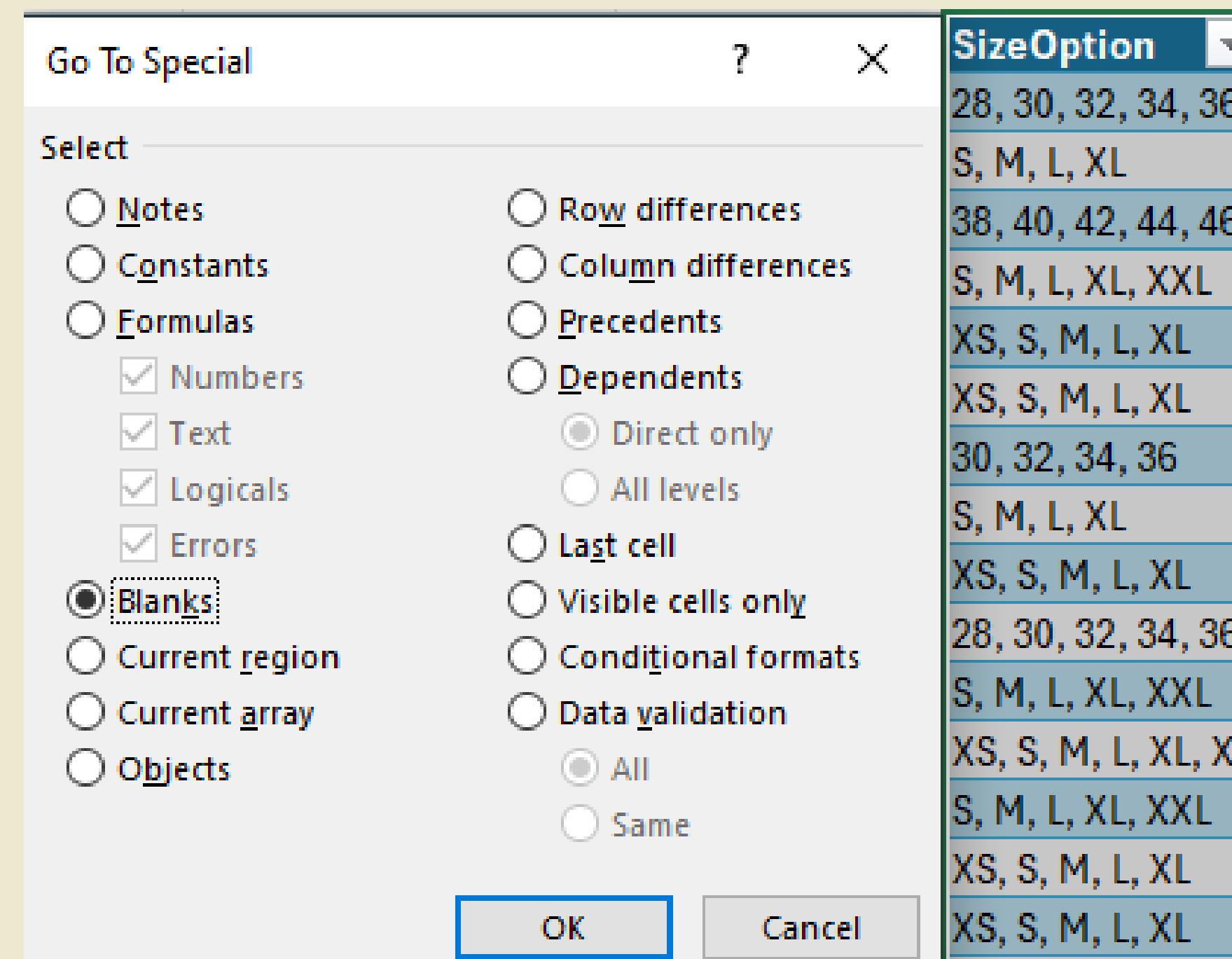
After getting discount value in single format in “Rs”. We will apply the formula

=IF(AND([@DiscountPrice (in Rs)]=" ",[@Discounter]=" "),AVERAGEIF([Category],[@Category],[@DiscountPrice (in Rs)]),[@DiscountPrice (in Rs)])

=IF(AND([@DiscountPrice (in Rs)]=" ",[@Discounter]=" "),AVERAGEIF([Category],[@Category],[@DiscountPrice (in Rs)]),[@DiscountPrice (in Rs)])

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

- a. Select the Size column
- b. press “Ctrl+G, select” Special”, click on” Blanks”
- c. All Blanks will be highlighted
- d. Type “ Not Available” and press Ctrl+Enter to fill all the selected Blanks



## B. 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."

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

Use the “ AVERAGEIF” functions as follows :

=AVERAGEIF(L:L, “>4”,I:I)

| =AVERAGEIF(L:L, “>4”,I:I)                               |                       |                       |            |                      |         |  |
|---|-----------------------|-----------------------|------------|----------------------|---------|--|
| G   | H                     | I                     | J          | K                    | L       |  |
| Description   | DiscountPrice (in Rs) | OriginalPrice (in Rs) | Discounter | SizeOption           | Ratings |  |
| roadster men n  | 824                   | 1499                  | 45%        | 28, 30, 32, 34, 36   | 3.9     |  |
| locomotive men  | 517                   | 1149                  | 55%        | S, M, L, XL          | 4       |  |
| roadster men n  | 629                   | 1399                  | 55%        | 38, 40, 42, 44, 46,  | 4.3     |  |
| zivame women  | 893                   | 1295                  | 31%        | S, M, L, XL, XXL     | 4.2     |  |
| roadster women white solid v neck pure cotton t shirt   |                       | 599                   | 35%        | XS, S, M, L, XL      | 4.2     |  |
| mast harbour women yellow solid tank top                |                       | 599                   | 40%        | XS, S, M, L, XL      | 4.4     |  |
| highlander men  | 599                   | 1499                  | 60%        | 30, 32, 34, 36       | 3.9     |  |
| mayra pink embroidered a line pure cotton t shirt       |                       | 1395                  | 58%        | S, M, L, XL          | 3.7     |  |
| roadster women pack of 2 solid t shirts                 |                       | 1098                  |            | XS, S, M, L, XL      | 4.3     |  |
| herenow men blue slim fit mid rise clean cotton t shirt |                       | 2749                  |            | 28, 30, 32, 34, 36   | 3.5     |  |
| hrx by hrithik rc                                       | 1214                  | 2699                  | 55%        | S, M, L, XL, XXL     | 4.4     |  |
| roadster men navy blue white striped polo shirt         |                       | 699                   |            | XS, S, M, L, XL, XXL | 4.1     |  |
| anubhutee women   | 1019                  | 3399                  | 70%        | S, M, L, XL, XXL     | 4.2     |  |
| athena women black solid basic jumpsuit                 |                       | 2499                  | 50%        | XS, S, M, L, XL      | 4.3     |  |
| roadster women maroon solid round neck t shirt          |                       | 799                   | 60%        | XS, S, M, L, XL      | 4       |  |

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

Step 1: Standardize discount format in percentage: =[@[Discounted price in Rs.]]/[@OriginalPrice (in Rs)]

Step 2: Use the formula to Count product with discount greater than 50%

| J         | K                   | L       | M       | N       | O        | P                  | Q |
|-----------|---------------------|---------|---------|---------|----------|--------------------|---|
| Discouter | SizeOption          | Ratings | Reviews | Column  | Column   |                    |   |
| 45%       | 28, 30, 32, 34, 36  | 3.9     | 999     | 674.55  | 1589.896 |                    |   |
| 55%       | S, M, L, XL         | 4       | 999     | 631.95  | 1589.896 | =COUNTIF(J:J,>50%) |   |
| 55%       | 38, 40, 42, 44, 46, | 4.3     | 999     | 769.45  | 1589.896 |                    |   |
| 31%       | S, M, L, XL, XXL    | 4.2     | 999     | 401.45  | 1589.896 |                    |   |
| 35%       | XS, S, M, L, XL     | 4.2     | 999     | 209.65  | 1589.896 |                    |   |
| 40%       | XS, S, M, L, XL     | 4.4     | 999     | 239.6   | 1589.896 |                    |   |
| 60%       | 30, 32, 34, 36      | 3.9     | 998     | 899.4   | 1589.896 |                    |   |
| 58%       | S, M, L, XL         | 3.7     | 998     | 809.1   | 1589.896 |                    |   |
|           | XS, S, M, L, XL     | 4.3     | 997     | 0       | 1589.896 |                    |   |
|           | 28, 30, 32, 34, 36  | 3.5     | 996     | 0       | 1589.896 |                    |   |
| 55%       | S, M, L, XL, XXL    | 4.4     | 996     | 1484.45 | 1589.896 |                    |   |

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

Use the formula : =COUNTIF(K:K,"\*M")

| K                   | L       | M       | N       | O        | P      | Q | R                  | S |
|---------------------|---------|---------|---------|----------|--------|---|--------------------|---|
| SizeOption          | Ratings | Reviews | Column  | Column   | Column |   |                    |   |
| 28, 30, 32, 34, 36  | 3.9     | 999     | 674.55  | 1589.896 | 519    |   |                    |   |
| S, M, L, XL         | 4       | 999     | 631.95  | 1589.896 | 519    |   |                    |   |
| 38, 40, 42, 44, 46, | 4.3     | 999     | 769.45  | 1589.896 | 519    |   | =COUNTIF(K:K,"*M") |   |
| S, M, L, XL, XXL    | 4.2     | 999     | 401.45  | 1589.896 | 519    |   |                    |   |
| XS, S, M, L, XL     | 4.2     | 999     | 209.65  | 1589.896 | 519    |   |                    |   |
| XS, S, M, L, XL     | 4.4     | 999     | 239.6   | 1589.896 | 519    |   |                    |   |
| 30, 32, 34, 36      | 3.9     | 998     | 899.4   | 1589.896 | 519    |   |                    |   |
| S, M, L, XL         | 3.7     | 998     | 809.1   | 1589.896 | 519    |   |                    |   |
| XS, S, M, L, XL     | 4.3     | 997     | 0       | 1589.896 | 519    |   |                    |   |
| 28, 30, 32, 34, 36  | 3.5     | 996     | 0       | 1589.896 | 519    |   |                    |   |
| S, M, L, XL, XXL    | 4.4     | 996     | 1484.45 | 1589.896 | 519    |   |                    |   |
| XS, S, M, L, XL, XX | 4.1     | 996     | 0       | 1589.896 | 519    |   |                    |   |
| S, M, L, XL, XXL    | 4.2     | 996     | 2379.3  | 1589.896 | 519    |   |                    |   |

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."

Step 1: Standardize discount format in percentage: =[@[Discounted price in Rs.]]/[@[OriginalPrice (in Rs)]]  
(already performed)

Step 2: Use formula = IF([@Discounted price in percentage]>50%, "high discount", "low discount")

| M       | N       | O        | P      | Q                              | R             | S |
|---------|---------|----------|--------|--------------------------------|---------------|---|
| Reviews | Column  | Column   | Column | Discounted price in percentage |               |   |
| 999     | 674.55  | 1589.896 | 519    | 55%                            | high discount |   |
| 999     | 631.95  | 1589.896 | 519    | 45%                            | low discount  |   |
| 999     | 769.45  | 1589.896 | 519    | 45%                            | low discount  |   |
| 999     | 401.45  | 1589.896 | 519    | 69%                            | high discount |   |
| 999     | 209.65  | 1589.896 | 519    | 0%                             | low discount  |   |
| 999     | 239.6   | 1589.896 | 519    | 0%                             | low discount  |   |
| 998     | 899.4   | 1589.896 | 519    | 40%                            | low discount  |   |
| 998     | 809.1   | 1589.896 | 519    | 0%                             | low discount  |   |
| 997     | 0       | 1589.896 | 519    | 0%                             | low discount  |   |
| 996     | 0       | 1589.896 | 519    | 0%                             | low discount  |   |
| 996     | 1484.45 | 1589.896 | 519    | 45%                            |               |   |

### C. Data Retrieval and Lookup

Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product\_id "2296012".

Find the "DiscountPrice" for the product with the Product ID "11895958" using the INDEX and MATCH functions.

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

1. Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product\_id "2296012"

```
=VLOOKUP(V2,Table1[[Product_id]:[Discounted price in percentage]],2,0)
```

```
=VLOOKUP(V2,Table1[[Product_id]:[Discounted price in percentage]],8,0)
```

```
=VLOOKUP(V2,Table1[[Product_id]:[Discounted price in percentage]],11,0)
```

```
=XLOOKUP(V2,Table1[Product_id],Table1[BrandName])
```

```
=XLOOKUP(V2,Table1[Product_id],Table1[OriginalPrice (in Rs)])
```

```
=XLOOKUP(V2,Table1[Product_id],Table1[Ratings])
```

2. Find the "DiscountPrice" for the product with the Product ID "11895958" using the INDEX and MATCH functions.

```
=INDEX(Table1[DiscountPrice (in Rs)],MATCH(Y3,Table1[Product_id],0))
```

```
=INDEX(Table1[DiscountPrice (in Rs)],MATCH(U3,Table1[Product_id],0))
```

| O        | P      | Q                              | R | S | T | U        |
|----------|--------|--------------------------------|---|---|---|----------|
| Column   | Column | Discounted price in percentage |   |   |   |          |
| L589.896 | 519    | 55%                            |   |   |   |          |
| L589.896 | 519    | 45%                            |   |   |   | 11895958 |
| L589.896 | 519    | 45%                            |   |   |   |          |
| L589.896 | 519    | 69%                            |   |   |   | 629      |
| L589.896 | 519    | 0%                             |   |   |   |          |
| L589.896 | 519    | 0%                             |   |   |   |          |
| L589.896 | 519    | 40%                            |   |   |   |          |

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

Step 1: Use the Data validation option to create a drop\_down for Product ID and details of products

Step 2: Use the Formula

| <code>=XLOOKUP(S2,Table1[Product_id],XLOOKUP(S5,Table1[#Headers],Table1))</code> |     |          |   |   |   |
|--|-----|----------|---|---|---|
| Q  | R   | S        | T | U | V |
| Discounted price in percentage   |     |          |   |   |   |
|  | 55% | 11895958 |   |   |   |
|  | 45% |          |   |   |   |
|  | 45% |          |   |   |   |
|  | 69% | Category |   |   |   |
|  | 0%  |          |   |   |   |
|  | 0%  | Topwear  |   |   |   |
|  | 40% |          |   |   |   |
|  | 0%  |          |   |   |   |
|  | 0%  |          |   |   |   |
|  | 0%  |          |   |   |   |

# SUMMARY:



This project focuses on analyzing Myntra's apparel dataset to gain insights into pricing, discounts, ratings, and product availability. The process involves three main stages:

## Data Cleaning & Preparation

Removing duplicate values.

Standardizing the discount column for uniformity.

Handling missing values in discount prices and size options.

## Data Analysis

Calculating average original price for highly rated products.

Identifying products with high discounts (>50%).

Counting availability by size (e.g., "M").

Categorizing products into "High Discount" and "Low Discount."

## Data Retrieval & Lookup

Using Excel functions like VLOOKUP, XLOOKUP, INDEX, and MATCH to find product details by product ID.

In short: The project demonstrates how data cleaning, analysis, and Excel lookup functions can be applied to a real-world e-commerce dataset, helping Myntra's management make better decisions about pricing, discounts, and inventory.

Do you also want me to create a Key Findings + Recommendations section so it looks more professional for your submission?

# KEY INSIGHTS:

Many products had inconsistent or missing values, highlighting the need for proper data cleaning.

A significant number of products were offering high discounts (>50%), showing strong price competition.

Products with ratings above 4 generally had higher average prices, reflecting customer trust in quality.

Certain sizes, especially “M”, showed higher availability, indicating higher demand or stocking preference.

Categorization into High Discount vs. Low Discount products helps identify potential sales strategies.

Excel functions like VLOOKUP/XLOOKUP enabled quick retrieval of product details, improving analysis efficiency.



# RECOMMENDATIONS:

Improve Data Quality – Regularly clean and standardize data to avoid missing or inconsistent values.

Optimize Discount Strategy – Balance high discounts with profitability by targeting discounts to slow-moving items.

Leverage Customer Ratings – Promote products with ratings above 4 to build trust and boost sales.

Enhance Inventory Planning – Monitor demand for popular sizes like “M” to prevent stockouts.

Segment Products – Use High vs. Low Discount categories to design personalized marketing campaigns.

Use Advanced Analytics – Apply Excel/SQL automation for faster insights into pricing, discounts, and availability trends.



# THANKU SO MUCH

*Presentation by: VIKASH SAINI*

*EMAIL: Vikashsaini2810@gmail.com*

*Contact no: 9784334174*

*Linkedin: <https://www.linkedin.com/in/vikash-saini-0366192a8/>*