# Analysis of Myntra Apparel

# About Company

Myntra is a leading Indian e-commerce platform specializing in fashion, lifestyle, and beauty products. Established in 2007, it has grown to become a one-stop destination for trend-conscious shoppers, offering a vast selection of clothing, footwear, accessories, and cosmetics from over 5,000 national and international brands. Myntra stands out for its focus on personalization, seamless user experience, and exclusive brand launches. As a subsidiary of the Flipkart Group (a Walmart-owned company), Myntra has played a pivotal role in shaping India's online fashion industry. With innovations like AI-driven styling, try-on features, and influencer-led campaigns, Myntra continues to redefine how India shops for fashion online.

# Project Questions

**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."
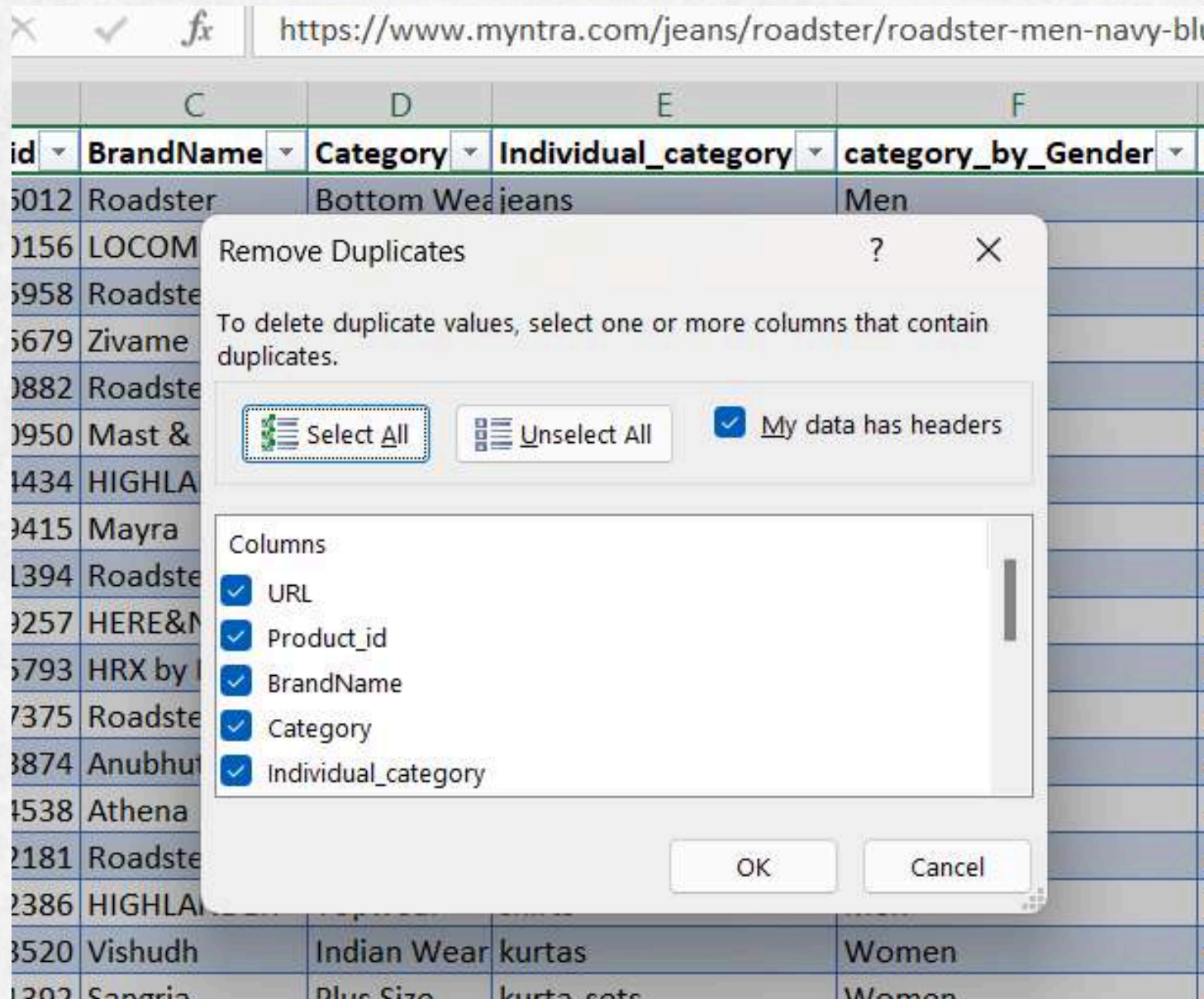
**B. Data Analysis**
1. Calculate the overall average original price for products with ratings greater than 4.
2. Count the number of products with a discount offer greater than 50% OFF.
3. Count the number of products available in size "M."
4. Create a new column to label the products as "High Discount" 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.

# A. Data Cleaning & Preparation

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



No Duplicates found in the given data.

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



Will first find and replace ",Hurry*" from "DiscountOffer " Coulum.

N2    =IF(ISNUMBER(SEARCH("Rs.",[@DiscounteOffer])),[@DiscounteOffer],[@DiscounteOffer]*[@[OriginalPrice (in Rs)]])

| | OriginalPrice (in Rs) | DiscounteOffer | SizeOption | Ratings | Reviews | Standardised DiscountOffer | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | 1499 | 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | | | |
| 3 | 1149 | 55% | S, M, L, XL | 4 | 999 | 631.95 | | | |
| 4 | 1399 | 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | | | |
| 5 | 1295 | 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | | | |
| 6 | 599 | 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | | | |
| 7 | 599 | 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | | | |
| 8 | 1499 | 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | | | |
| 9 | 1395 | 58% | S, M, L, XL | 3.7 | 998 | 809.10 | | | |
| 10 | 1098 | | XS, S, M, L, XL | 4.3 | 997 | 0.00 | | | |
| 11 | 2749 | | 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | | | |
| 12 | 2699 | 55% | S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | | | |
| 13 | 699 | | XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | | | |
| 14 | 3399 | 70% | S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | | | |
| 15 | 2499 | 50% | XS, S, M, L, XL | 4.3 | 996 | 1249.50 | | | |
| 16 | 799 | 60% | XS, S, M, L, XL | 4 | 996 | 479.40 | | | |
| 17 | 1099 | 53% | 39, 40, 42, 44 | 4.2 | 995 | 582.47 | | | |
| 18 | 1699 | 59% | S, M, L, XL, XXL | 4.2 | 995 | 1002.41 | | | |
| 19 | 3999 | | XS, S, M, L, XL, XXL, 3XL | 4.3 | 995 | 0.00 | | | |
| 20 | 2149 | 61% | 28, 30, 32, 34 | 4.1 | 995 | 1310.89 | | | |

Then will put this formula and
remove Rs. to standardise values

# 3. Identify 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 | | Cells | |

**P3** | =IF(AND(ISBLANK([@[DiscountPrice (in Rs)]]), ISBLANK([@[Standardised DiscountOffer]])), AVERAGEIFS(H:H, D:D, [@Category]), [@[DiscountPrice (in Rs)]])

| | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | DiscountPrice (in Rs) | OriginalPrice (in I | DiscounteOffe | SizeOption | Ratings | Reviews | Standardised DiscountOff | Column1 | Columr | | |
| 2 | 824 | 1499 | 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | | 824 | | |
| 3 | 517 | 1149 | 55% | S, M, L, XL | 4 | 999 | 631.95 | | 517 | | |
| 4 | 629 | 1399 | 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | | 629 | | |
| 5 | 893 | 1295 | 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | | 893 | | |
| 6 | | 599 | 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | | 0 | | |
| 7 | | 599 | 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | | 0 | | |
| 8 | 599 | 1499 | 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | | 599 | | |
| 9 | | 1395 | 58% | S, M, L, XL | 3.7 | 998 | 809.10 | | 0 | | |
| 10 | | 1098 | | XS, S, M, L, XL | 4.3 | 997 | 0.00 | Not Available | 0 | | |
| 11 | | 2749 | | 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | Not Available | 0 | | |
| 12 | 1214 | 2699 | 55% | S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | | 1214 | | |
| 13 | | 699 | | XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | Not Available | 0 | | |
| 14 | 1019 | 3399 | 70% | S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | | 1019 | | |
| 15 | | 2499 | 50% | XS, S, M, L, XL | 4.3 | 996 | 1249.50 | | 0 | | |
| 16 | | 799 | 60% | XS, S, M, L, XL | 4 | 996 | 479.40 | | 0 | | |
| 17 | 516 | 1099 | 53% | 39, 40, 42, 44 | 4.2 | 995 | 582.47 | | 516 | | |
| 18 | 696 | 1699 | 59% | S, M, L, XL, XXL | 4.2 | 995 | 1002.41 | | 696 | | |
| 19 | | 3999 | | XS, S, M, L, XL, XXL, 3XL | 4.3 | 995 | 0.00 | Not Available | 0 | | |
| 20 | | 2149 | 61% | 28, 30, 32, 34 | 4.1 | 995 | 1310.89 | | 0 | | |
| 21 | | 999 | 75% | XS, S, M, L, XL | 4.3 | 995 | 749.25 | | 0 | | |
| 22 | | 1298 | | XS, S, M, L, XL | 4.4 | 994 | 0.00 | Not Available | 0 | | |

Using the formula
"=IF(AND(ISBLANK([@[DiscountPrice (in Rs)]]), ISBLANK([@[Standardised DiscountOffer]])), AVERAGEIFS(H:H, D:D, [@Category]), [@[DiscountPrice (in Rs)]])"
we can solve this query and get the results in column P.

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

| rice (in Rs) | OriginalPrice (in I | DiscounteOffe | SizeOption | Ratings | Reviews | Standardised DiscountOff | Column1 | Colum |
|---|---|---|---|---|---|---|---|---|
| 1499 | 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | | | 824 |
| 1149 | 55% | S, M, L, XL | 4 | 999 | 631.95 | | | 517 |
| 1399 | 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | | | 629 |
| 1295 | 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | | | 893 |
| 599 | 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | | | 0 |
| 599 | 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | | | 0 |
| 1499 | 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | | | 599 |
| 1395 | 58% | S, M, L, XL | | | | | | 0 |
| 1098 | | XS, S, M, L, XL | | | | | | 0 |
| 2749 | | 28, 30, 32, 34, 36 | | | | | | 0 |
| 2699 | 55% | S, M, L, XL, XXL | | | | | | 1214 |
| 699 | | XS, S, M, L, XL, XXL, 3XL, 4X | | | | | | 0 |
| | | | | | | | | 1019 |
| | | | | | 1249.50 | | | 0 |
| | | | | | 479.40 | | | 0 |
| | | | | | 582.47 | | | 516 |
| | | | | | 1002.41 | | | 696 |
| | | | | | 0.00 | | Not Available | 0 |
| | | | | | 1310.89 | | | 0 |
| | | | | | 749.25 | | | 0 |
| | | | | | 0.00 | | Not Available | 0 |
| | | | | | 879.45 | | | 719 |
| | | | | | 0.00 | | Not Available | 0 |
| | | | | | 749.50 | | | 0 |
| | | | | | 805.38 | | | 0 |
| 1299 | 57% | S, M, L, XL | 4.2 | 991 | 740.43 | | | 0 |
| 2699 | 70% | 26, 28, 30, 32, 34, 36 | 4.1 | 991 | 1889.30 | | | 0 |

**Microsoft Excel** ✕

⚠ We couldn't find what you were looking for. Click Options for more ways to search.

[ OK ]

## Find and Replace

**Find** | **Replace**

Find what: | Null | | No Format Set | Format... ▾

☐ Match case
☐ Match entire cell contents

Within: Sheet
Search: By Rows
Look in: Formulas

Options <<

[ Find All ] [ Find Next ] [ Close ]

No null value found in the "SizeOption" section.

# B. Data Analysis

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



| | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|
| 1 | OriginalPrice (in R ▾ | DiscounteOffe ▾ | SizeOption ▾ | Ratings ▾ | Reviews ▾ | Standardised DiscountOffe ▾ | Column1 ▾ | Columr ▾ | Columr ▾ |
| 2 | 1499 | 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | | 824 | |
| 3 | 1149 | 55% | S, M, L, XL | 4 | 999 | 631.95 | | 517 | |
| 4 | 1399 | 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | | 629 | |
| 5 | 1295 | 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | | 893 | |
| 6 | 599 | 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | | 0 | |
| 7 | 599 | 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | | 0 | |
| 8 | 1499 | 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | | 599 | |
| 9 | 1395 | 58% | S, M, L, XL | 3.7 | 998 | 809.10 | | 0 | |
| 10 | 1098 | | XS, S, M, L, XL | 4.3 | 997 | 0.00 | Not Available | 0 | |
| 11 | 2749 | | 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | Not Available | 0 | |
| 12 | 2699 | 55% | S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | | 1214 | |
| 13 | 699 | | XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | Not Available | 0 | |
| 14 | 3399 | 70% | S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | | 1019 | |
| 15 | 2499 | 50% | XS, S, M, L, XL | 4.3 | 996 | 1249.50 | | 0 | |
| 16 | 799 | 60% | XS, S, M, L, XL | 4 | 996 | 479.40 | | 0 | |
| 17 | 1099 | 53% | 39, 40, 42, 44 | 4.2 | 995 | 582.47 | | 516 | |
| 18 | 1699 | 59% | S, M, L, XL, XXL | 4.2 | 995 | 1002.41 | | 696 | |
| 19 | 3999 | | XS, S, M, L, XL, XXL, 3XL | 4.3 | 995 | 0.00 | Not Available | 0 | |
| 20 | 2149 | 61% | 28, 30, 32, 34 | 4.1 | 995 | 1310.89 | | 0 | |
| 21 | 999 | 75% | XS, S, M, L, XL | 4.3 | 995 | 749.25 | | 0 | |
| 22 | 1298 | | XS, S, M, L, XL | 4.4 | 994 | 0.00 | Not Available | 0 | |
| 23 | 1599 | 55% | XS, S, M, L, XL, XXL | 3.8 | 993 | 879.45 | | 719 | |
| 24 | 849 | | S, M, L, XL | 4.2 | 993 | 0.00 | Not Available | 0 | |
| 25 | 1499 | 50% | XS, S, M, L, XL | 4.2 | 993 | 749.50 | | 0 | |
| 26 | 1299 | 62% | XS, S, M, L, XL, XXL | 4.3 | 992 | 805.38 | | 0 | |
| 27 | 1299 | 57% | S, M, L, XL | 4.2 | 991 | 740.43 | | 0 | |
| 28 | 2699 | 70% | 26, 28, 30, 32, 34, 36 | 4.1 | 991 | 1889.30 | | 0 | |
| 29 | 899 | 61% | 26, 28, 30, 32, 34, 36 | 4 | 991 | 548.39 | | 0 | |
| 30 | 899 | 60% | XS, S, M, L, XL, XXL, 3XL | 4 | 990 | 539.40 | | 359 | |

=AVERAGEIF(L:L, ">4", I:I)
AVERAGEIF(range, criteria, [average_range])

In this will find the value using the formula " =AVERAGEIF(L:L, ">4", I:I) " from column "I" and "L".

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



`=COUNTIF(J:J, ">50%")`

| DiscounteOffer | SizeOption | Ratings | Reviews | Standardised DiscountOf | Colum |
|---|---|---|---|---|---|
| 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | |
| 55% | S, M, L, XL | 4 | 999 | 631.95 | |
| 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | |
| 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | |
| 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | |
| 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | |
| 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | |
| 58% | S, M, L, XL | 3.7 | 998 | 809.10 | |
| | XS, S, M, L, XL | 4.3 | 997 | 0.00 | |
| | 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | |
| 55% | S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | |
| | XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | |
| 70% | S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | |
| 50% | XS, S, M, L, XL | 4.3 | 996 | 1249.50 | |
| 60% | XS, S, M, L, XL | 4 | 996 | 479.40 | |
| 53% | 39, 40, 42, 44 | 4.2 | 995 | 582.47 | |
| 59% | S, M, L, XL, XXL | 4.2 | 995 | 1002.41 | |
| | XS, S, M, L, XL, XXL, 3XL | 4.3 | 995 | 0.00 | |
| 61% | 28, 30, 32, 34 | 4.1 | 995 | 1310.89 | |
| 75% | XS, S, M, L, XL | 4.3 | 995 | 749.25 | |
| | XS, S, M, L, XL | 4.4 | 994 | 0.00 | |
| 55% | XS, S, M, L, XL, XXL | 3.8 | 993 | 879.45 | |
| | S, M, L, XL | 4.2 | 993 | 0.00 | |
| 50% | XS, S, M, L, XL | 4.2 | 993 | 749.50 | |
| 62% | XS, S, M, L, XL, XXL | 4.3 | 992 | 805.38 | |
| 57% | S, M, L, XL | 4.2 | 991 | 740.43 | |
| 70% | 26, 28, 30, 32, 34, 36 | 4.1 | 991 | 1889.30 | |
| 61% | 26, 28, 30, 32, 34, 36 | 4 | 991 | 548.39 | |
| 60% | XS, S, M, L, XL, XXL, 3XL | 4 | 990 | 539.40 | |
| 55% | 26, 28, 30, 32, 34, 36 | 4.3 | 990 | 659.45 | |
| 50% | L, M, S, XS, XL | 4.2 | 990 | 799.00 | |
| 65% | S, M, L, XL, XXL | 4.4 | 990 | 1526.85 | |

`=COUNTIF(J:J, ">50%")`

Using the formula "=COUNTIF(J:J, ">50%") "will find the number of products.

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

| SizeOption | Ratings | Reviews | Standardised DiscountOf | Column |
|---|---|---|---|---|
| 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | |
| S, M, L, XL | 4 | 999 | 631.95 | |
| 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | |
| S, M, L, XL, XXL | 4.2 | 999 | 401.45 | |
| XS, S, M, L, XL | 4.2 | 999 | 209.65 | |
| XS, S, M, L, XL | 4.4 | 999 | 239.60 | |
| 30, 32, 34, 36 | 3.9 | 998 | 899.40 | |
| S, M, L, XL | 3.7 | 998 | 809.10 | |
| XS, S, M, L, XL | 4.3 | 997 | 0.00 | |
| 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | |
| S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | |
| XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | |
| S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | |
| XS, S, M, L, XL | 4.3 | 996 | 1249.50 | |
| XS, S, M, L, XL | 4 | 996 | 479.40 | |
| 39, 40, 42, 44 | 4.2 | 995 | 582.47 | |
| S, M, L, XL, XXL | 4.2 | 995 | 1002.41 | |
| XS, S, M, L, XL, XXL, 3XL | 4.3 | 995 | 0.00 | |
| 28, 30, 32, 34 | 4.1 | 995 | 1310.89 | |
| XS, S, M, L, XL | 4.3 | 995 | 749.25 | |
| XS, S, M, L, XL | 4.4 | 994 | 0.00 | |
| XS, S, M, L, XL, XXL | 3.8 | 993 | 879.45 | |
| S, M, L, XL | 4.2 | 993 | 0.00 | |
| XS, S, M, L, XL | 4.2 | 993 | 749.50 | |
| XS, S, M, L, XL, XXL | 4.3 | 992 | 805.38 | |
| S, M, L, XL | 4.2 | 991 | 740.43 | |
| 26, 28, 30, 32, 34, 36 | 4.1 | 991 | 1889.30 | |
| 26, 28, 30, 32, 34, 36 | 4 | 991 | 548.39 | |
| XS, S, M, L, XL, XXL, 3XL | 4 | 990 | 539.40 | |
| 26, 28, 30, 32, 34, 36 | 4.3 | 990 | 659.45 | |
| L, M, S, XS, XL | 4.2 | 990 | 799.00 | |
| S, M, L, XL, XXL | 4.4 | 990 | 1526.85 | |

=COUNTIF(K:K, "*M*")

In this also we use "COUNTIF" function.
As we have to find all products available in size "M" not only "M", so will put "*" both the side of it "*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."

| | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|
| | DiscounteOffer ▾ | SizeOption ▾ | Ratings ▾ | Reviews ▾ | Standardised DiscountOf ▾ | Colum ▾ | | | |
| | 45% | 28, 30, 32, 34, 36 | 3.9 | 999 | 674.55 | =IF([@DiscounteOffer] >50%, "High | | | |
| | 55% | S, M, L, XL | 4 | 999 | 631.95 | Discount", "Low Discount") | | | |
| | 55% | 38, 40, 42, 44, 46, 48 | 4.3 | 999 | 769.45 | IF(**logical_test**, [value_if_true], [value_if_false]) | | | |
| | 31% | S, M, L, XL, XXL | 4.2 | 999 | 401.45 | Low Discount | | | |
| | 35% | XS, S, M, L, XL | 4.2 | 999 | 209.65 | Low Discount | | | |
| | 40% | XS, S, M, L, XL | 4.4 | 999 | 239.60 | Low Discount | | | |
| | 60% | 30, 32, 34, 36 | 3.9 | 998 | 899.40 | High Discount | | | |
| | 58% | S, M, L, XL | 3.7 | 998 | 809.10 | High Discount | | | |
| | | XS, S, M, L, XL | 4.3 | 997 | 0.00 | Low Discount | | | |
| | | 28, 30, 32, 34, 36 | 3.5 | 996 | 0.00 | Low Discount | | | |
| | 55% | S, M, L, XL, XXL | 4.4 | 996 | 1484.45 | High Discount | | | |
| | | XS, S, M, L, XL, XXL, 3XL, 4XL | 4.1 | 996 | 0.00 | Low Discount | | | |
| | 70% | S, M, L, XL, XXL | 4.2 | 996 | 2379.30 | High Discount | | | |

Using the formula "=IF([@DiscounteOffer] >50%, "High Discount", "Low Discount")", will get the result from "DiscountOffer".

# C. Data Retrieval and Lookup

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

```
=XLOOKUP(11226634,
B:B,C:C, 0 )
```

"C" is the column containing "Brand Name".

```
=XLOOKUP(11226634,
B:B,I:I, 0 )
```

"I" is the column containing "Original Price".

```
=XLOOKUP(11226634,
B:B,L:L, 0 )
```

"L" is the column containing "Ratings".

11226634

=VLOOKUP(P10, B:M, 2,FALSE)

11226634

=VLOOKUP(P10, B:M, 8,FALSE)

11226634

=VLOOKUP(P10, B:M, 11,FALSE)

In this the range of table starts from the reference column.

"2" is the "Brand Name" column.
"8" is the "Original Price" column.
"11" is the "Ratings" column.

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

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

INDEX(array, row_num, [column_num])
INDEX(reference, row_num, [column_num], [area_num])

Using this formula will get the "DiscountPrice" 599.