



# Myntra Apparel

## Data Analysis and Insights

**About the Company:** Myntra is a fashion-focused e-commerce company based in India offering apparel, lifestyle, and personal care brands.

## 1. Data Pre-processing

cleaning and standardization

## 2. Data Analysis

## 3. Data Retrieval and Lookup

# Data Pre-processing

## Client-Informed Data Issues and Applied Fixes:

- 1. Resolving data redundancy:** After converting data to table, check from duplicates at table level and in the primary key column.  
**For Duplicate rows in table:** Data Tab > Data Tools > Remove Duplicates > Select all  
**For Duplicates in primary key column:** Home Tab > Styles > Conditional Formatting > Highlight cells > Duplicate values, Sort and filter > Custom Sort > Sort by Primary key column and sort on cell color to highlight duplicates
- 2. Ensuring data consistency:** It was informed that the discount offer column is inconsistent, the following process was employed to standardize the column and ensure consistent values.

- Different types of discrepancies were found out using unique function to narrow down scope

```
=UNIQUE (
  Table1[DiscountOffer]
)
```

70% OFF	70%
61% OFF	61%
60% OFF, Hurry*	60%
55% OFF	55%
Rs. 500 OFF	500
61% OFF	61%

1

- Values were trimmed, applied with nested substitutes and trimmed again for further processing

```
=TRIM(
  SUBSTITUTE(
    SUBSTITUTE(
      TRIM(
        [@DiscountOffer]
      ),
      ", Hurry*", ""
    ),
    "Rs. ", ""
  ),
  "OFF", ""
)
```

2

- Conditionals used to find and convert percentages to discount offer values

```
=IF(
  ISNUMBER(
    FIND(
      "%",
      [@[Disc_offer_cleaned]]
    )
  ), LEFT(
    [@[Disc_offer_cleaned]],
    SEARCH(
      "%",
      [@[Disc_offer_cleaned]]
    ) - 1
  ) / 100 * [@[OriginalPrice (in Rs)]],
  [@[Disc_offer_cleaned]]
)
```

3

# Data Pre-processing

**3. Applying Category-aware data imputation:** Discount price column had blanks, they were replaced using discount values calculated in the previous step by the following process:

- Discount values were converted back to percentages and the data type was changed to general to use for calculations.
- Since there were blanks in this column as well, iferror function was used to retain blanks for this step

```
=IFERROR(  
  ( [@Disc_values]  
  / [@[OriginalPrice] * 100,""  
)
```

1

- The blanks in discount percent column thus created were filled using average by categories:

```
=IF(  
  [@[Disc_Percent_w_nulls]] = "",  
  AVERAGEIFS(  
    [Disc_Percent_w_nulls],  
    [Category],  
    [@Category]  
  ),  
  [@[Disc_Percent_w_nulls]]  
)
```

2

- Then the following conditional was employed to fill blanks in discount price column and the values were rounded off.

```
=FLOOR.MATH(  
  IF(  
    [@[DiscountPrice_in_Rs]] = "",  
    [@[OriginalPrice (in Rs)]] -  
    ( [@[Discount_Percent_wo_nulls]]  
      / 100 ) * [@[OriginalPrice]],  
    [@[DiscountPrice_in_Rs]]  
  )  
)
```

3

**4. Handling Nulls:** Null values, if any in the sizes column were to be found and replaced by “Not Available”. But upon querying using the quick sort button on the table, no blanks were found.

# Data Pre-processing

---

## Report

### 1. Resolving data redundancy

- No duplicate rows were detected in the data
- Primary key column had no duplicates

### 2. Ensuring data consistency

- Discount offer column was standardized, it was decided to convert percentages to actual values to maintain consistency and feasibility for analysis

### 3. Applying Category-aware data imputation

- The blank cells in the discount price column were filled by first converting discount offer column to percentages and taking it's average as per product categories
- It was decided to use the average of percent rather than the actual discount offer values to avoid influence of large variation and outliers in the actual values. Percentages provided a better average estimate.
- Blank values were filled in the discount price column using conditionals.

### 4. Handling nulls

- Upon checking the sizes column it was discovered that there were no null values to be replaced by "Not available" text.



# **Myntra Apparel**

## **Data Analysis and Insights**

### 1. Data pre-processing

cleaning and standardization

### 2. Data Analysis

### 3. Data retrieval and lookup

# Data Analysis

## Client-Informed Analysis Requirements

1. Overall Average  
Original Price of High-  
Rated Products

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

**Rs. 1618**

2. Discount Depth  
Analysis: Products  
with >50% OFF

```
=COUNTIF(  
  Table1[Discount_Percent_wo  
_nulls],  
  ">50"  
)
```

**Count: 3308**

3. Inventory Availability:  
Products in Size 'M'

```
=COUNTA(  
  Table1[Column1]  
) -  
COUNTBLANK(  
  Table1[Column1]  
)
```

**Count: 3693**

4. Binary Classification  
of Discount Offers

```
=IF(  
  [@[Discount_Percent_wo_nu  
lls]] > 50,  
  "High Discount",  
  "Low Discount"  
)
```

**High Disc.: 3308  
Low Disc.: 2249**



# Myntra Apparel

## Data Analysis and Insights

### 1. Data pre-processing

cleaning and standardization

### 2. Data Analysis

### 3. Data Retrieval and Lookup

# Data Retrieval and Lookup

## Specific Retrieval requests:

1. Use 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.

```
=XLOOKUP(  
  11226634,  
  Table1[Product_id],  
  CHOOSE(  
    { 1,2,3},  
    Table1[BrandName],  
    Table1[OriginalPrice (in Rs)],  
    Table1[Ratings]  
  )  
)
```

**Maniac ,1199, 3.9**

```
=INDEX(  
  Table1[Discount_price],  
  MATCH(  
    6744434,  
    Table1[Product_id],  
    0  
  )  
)
```

**Rs. 599**

```
=XLOOKUP(  
  A118,  
  Table1[Product_id],  
  XLOOKUP(  
    AH18,  
    Table1[#Headers],  
    Table1  
  )  
)
```

- Create Data Validation Lists for both axes

BrandName ↓ 13205276 ↓  
'all about you'



# Data Retrieval and Lookup

## Appendix A

### 1. Reference screenshot from Excel workbook

1	URL	Product_id	BrandName	Category	Individual_category	category_b	Discou	OriginalP	Discount	Disc	Colou	Disc_P	Discount	SizeOption	Colu	Rating	Revi	Disc	Disc	Column	DiscountPrice
2	https://wv	2296012	Roadster	Bottom We	jeans	Men	824	1499	45% OFF	45%	674.55	45	45	28, 30, 32, 34,		3.9	999	824	824	Low Discou	bothNotNull
3	https://wv	13780156	LOCOMOTIVE	Bottom We	track-pants	Men	517	1149	55% OFF	55%	631.95	55	55	S, M, L, XL	4	4	999	517	517	High Discou	bothNotNull
4	https://wv	11895958	Roadster	Topwear	shirts	Men	629	1399	55% OFF	55%	769.45	55	55	38, 40, 42, 44,		4.3	999	629	629	High Discou	bothNotNull
5	https://wv	4335679	Zivame	Lingerie & Sl	shapewear	Women	893	1295	31% OFF	31%	401.45	31	31	S, M, L, XL, XXL	4	4.2	999	893	893	Low Discou	bothNotNull
6	https://wv	11690882	Roadster	Western	tshirts	Women		599	35% OFF	35%	209.65	35	35	XS, S, M, L, XL	8	4.2	999		389	Low Discou	bothNotNull
7	https://wv	2490950	Mast & Harbou	Western	tops	Women		599	40% OFF	40%	239.6	40	40	XS, S, M, L, XL	8	4.4	999		359	Low Discou	bothNotNull
8	https://wv	6744434	HIGHLANDER	Bottom We	trousers	Men	599	1499	60% OFF	60%	899.4	60	60	30, 32, 34, 36		3.9	998	599	599	High Discou	bothNotNull
9	https://wv	8439415	Mayra	Western	tops	Women		1395	58% OFF	58%	809.1	58	58	S, M, L, XL	4	3.7	998		585	High Discou	bothNotNull
10	https://wv	17381394	Roadster	Western	tshirts	Women		1098						XS, S, M, L, XL	8	4.3	997		516	High Discou	bothNull
11	https://wv	2359257	HERE&NOW	Bottom We	jeans	Men		2749						28, 30, 32, 34,		3.5	996		1284	High Discou	bothNull
12	https://wv	7695793	HRX by Hrithik F	Sports Wear	tights	Men	1214	2699	55% OFF	55%	1484.5	55	55	S, M, L, XL, XXL	4	4.4	996	1214	1214	High Discou	bothNotNull
13	https://wv	10307375	Roadster	Topwear	tshirts	Men		699						XS, S, M, L, XL,	8	4.1	996		328	High Discou	bothNull
14	https://wv	12873874	Anubhutee	Indian Wear	kurta-sets	Women	1019	3399	70% OFF	70%	2379.3	70	70	S, M, L, XL, XXL	4	4.2	996	1019	1019	High Discou	bothNotNull
15	https://wv	11634538	Athena	Western	jumpsuit	Women		2499	50% OFF	50%	1249.5	50	50	XS, S, M, L, XL	8	4.3	996		1249	Low Discou	bothNotNull
16	https://wv	2312181	Roadster	Western	tshirts	Women		799	60% OFF	60%	479.4	60	60	XS, S, M, L, XL	8	4	996		319	High Discou	bothNotNull
17	https://wv	13842386	HIGHLANDER	Topwear	shirts	Men	516	1099	53% OFF	53%	582.47	53	53	39, 40, 42, 44		4.2	995	516	516	High Discou	bothNotNull
18	https://wv	10473520	Vishudh	Indian Wear	kurtas	Women	696	1699	59% OFF	59%	1002.4	59	59	S, M, L, XL, XXL	4	4.2	995	696	696	High Discou	bothNotNull
19	https://wv	10561392	Sangria	Plus Size	kurta-sets	Women		3999						XS, S, M, L, XL,	8	4.3	995		1546	High Discou	bothNull
20	https://wv	12391750	Tokyo Talkies	Western	trousers	Women		2149	61% OFF	61%	1310.9	61	61	28, 30, 32, 34		4.1	995		838	High Discou	bothNotNull
21	https://wv	2522986	DressBerry	Western	tshirts	Women		999	75% OFF	75%	749.25	75	75	XS, S, M, L, XL	8	4.3	995		249	High Discou	bothNotNull
22	https://wv	17385142	Roadster	Western	tshirts	Women		1298						XS, S, M, L, XL	8	4.4	994		610	High Discou	bothNull
23	https://wv	12153330	Anouk	Indian Wear	kurtas	Women	719	1599	55% OFF	55%	879.45	55	55	XS, S, M, L, XL,	8	3.8	993	719	719	High Discou	bothNotNull
24	https://wv	1864573	Enamor	Lingerie & Sl	bra	Women		849						S, M, L, XL	4	4.2	993		475	Low Discou	bothNull
25	https://wv	13205276	all about you	Western	shirts	Women		1499	50% OFF	50%	749.5	50	50	XS, S, M, L, XL	8	4.2	993		749	Low Discou	bothNotNull
26	https://wv	11535928	KASSUALLY	Western	tops	Women		1299	62% OFF	62%	805.38	62	62	XS, S, M, L, XL,	8	4.3	992		493	High Discou	bothNotNull
27	https://wv	6552977	RARE	Western	tops	Women		1299	57% OFF	57%	740.43	57	57	S, M, L, XL	4	4.2	991		558	High Discou	bothNotNull