

Myntra Analytics Project

# Myntra Analytics Project

SHUBHAM SHARMA



### HELLO EVERYONE

#### I AM SHUBHAM SHARMA

As an aspiring data analyst, I applied my SQL, Python (Pandas, Matp<mark>lotlib), and Power BI</mark> Skills to analyze Myntra's sales and pricing data, uncovering key business insights to support data-driven decision-making.

Through my analysis, I explored how pricing strategies, discounts, customer satisfaction, brand performance, inventory flow, and category trends impact Myntra's business, uncovering insights that can guide smarter decisions.



### **Business Problem**

- Myntra, like any growing e-commerce platform, faces a few big questions. Are the discounts it gives really helping sales, or just cutting into profits? Are customers truly satisfied across categories, sizes, and price ranges—or are some areas letting them down? Some brands are thriving, while others survive only on heavy discounts—so how do we measure real brand strength? On the inventory side, Myntra often ends up with slow-moving products sitting unsold while fast-sellers risk running out of stock. And finally, which categories and customer segments—men, women, or specific styles—are actually driving the most engagement and growth? These are the problems I focused on solving through data.
- For this project, I started by exploring and cleaning the Myntra dataset using SQL. Then, with the help of Python libraries like Pandas, NumPy, and Matplotlib, I analyzed the data to uncover meaningful patterns. Finally, I brought everything together in Power BI, where I built an interactive dashboard to turn those numbers into easy-to-understand insights that connect directly to real business problems.



# Exploratory Data Analysis on Myntra Product Dataset

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read csv(
    r'C:\Users\ASUS\OneDrive\Desktop\Myntra Fasion Clothing.csv',
    dtype={9: str} )
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 526564 entries, 0 to 526563
Data columns (total 13 columns):
     Column
                            Non-Null Count
                                            Dtype
                            526564 non-null object
     Product id
                            526564 non-null int64
     BrandName
                            526564 non-null object
     Category
                            526564 non-null object
     Individual category
                            526564 non-null object
     category_by_Gender
                            526564 non-null object
     Description
                            526564 non-null object
     DiscountPrice (in Rs) 333406 non-null float64
     OriginalPrice (in Rs) 526564 non-null float64
     DiscountOffer
                            452258 non-null object
    SizeOption
                            526564 non-null object
 11 Ratings
                            190412 non-null float64
 12 Reviews
                            190412 non-null float64
dtypes: float64(4), int64(1), object(8)
memory usage: 52.2+ MB
```



# Myntra Exploratory Data Analysis

[4]:	df.head(10)								
[4]:	URL	Product_id	BrandName	Category	Individual_category	category_by_Gender	Description	DiscountPrice (in Rs)	OriginalPrice (in Rs)
	https://www.myntra.com/jeans/roadster/roadster	2296012	Roadster	Bottom Wear	jeans	Men	roadster men navy blue slim fit mid rise clean	824.0	1499.0
	https://www.myntra.com/track-pants/locomotive/	13780156	LOCOMOTIVE	Bottom Wear	track-pants	Men	locomotive men black white solid slim fit tra	517.0	1149.0
	https://www.myntra.com/shirts/roadster/roadste	11895958	Roadster	Topwear	shirts	Men	roadster men navy white black geometric print	629.0	1399.0
	https://www.myntra.com/shapewear/zivame/zivame	4335679	Zivame	Lingerie & Sleep Wear	shapewear	Women	zivame women black saree shapewear zi3023core0	893.0	1295.0
	https://www.myntra.com/tshirts/roadster/roadst	11690882	Roadster	Western	tshirts	Women	roadster women white solid v neck pure cotton 	NaN	599.0
							mast harbour		

```
df.isnull().sum()
[21]: URL
                                0
      Product_id
       BrandName
      Category
      Individual category
      category by Gender
      Description
      OriginalPrice (in Rs)
      DiscountOffer
      SizeOption
                                0
      Ratings
       Reviews
       DiscountPercent
      dtype: int64
      Filling Null Values
[23]: df['Ratings'] = df['Ratings'].fillna(df['Ratings'].mean())
      df['Reviews'] = df['Reviews'].fillna(df['Reviews'].mean())
      df['DiscountOffer'] = df['Reviews'].fillna(df['Reviews'].mean())
```

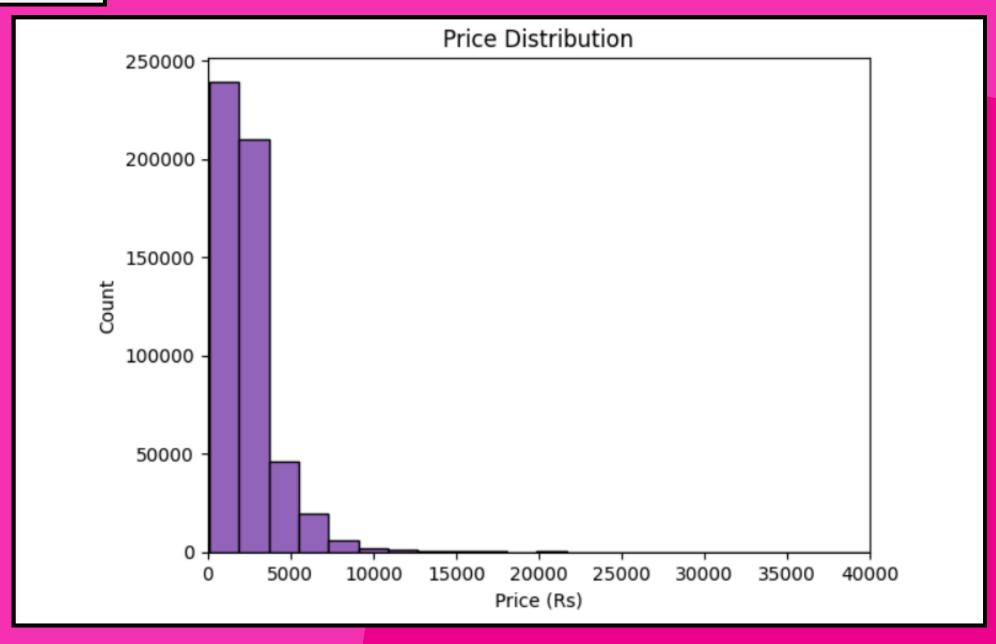
```
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 526564 entries, 0 to 526563
Data columns (total 12 columns):
     Column
                           Non-Null Count
                                           Dtype
    URL
                           526564 non-null object
     Product id
                           526564 non-null int64
     BrandName
                           526564 non-null object
                           526564 non-null object
     Category
    Individual_category
                           526564 non-null object
    category by Gender
                           526564 non-null object
                           526564 non-null object
    Description
    OriginalPrice (in Rs) 526564 non-null float64
    DiscountOffer
                           526564 non-null float64
    SizeOption
                           526564 non-null object
 10 Ratings
                           526564 non-null float64
 11 Reviews
                           526564 non-null float64
dtypes: float64(4), int64(1), object(7)
memory usage: 48.2+ MB
```

```
Price Distribution

•[19]: plt.hist(df['OriginalPrice (in Rs)'], bins=50, edgecolor='black')
    plt.xlabel("Price (Rs)")
    plt.ylabel("Count")
    plt.title("Price Distribution")
    plt.xlim(0, 40000)
    plt.show()
```

### Price Distribution Histogram





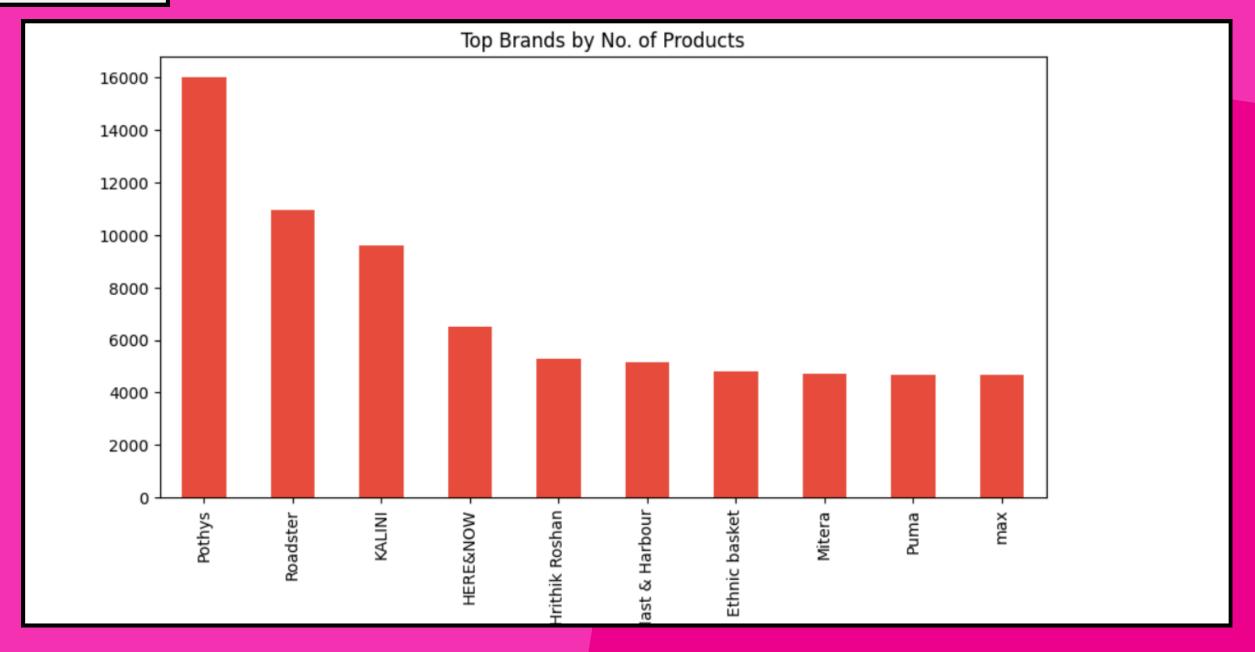


```
Do higher discounts lead to better ratings?
plt.figure(figsize=(6,4))
sns.heatmap(df[['DiscountPercent','Ratings']].corr(), annot=True, cmap='coolwarm', center=0)
plt.title("Correlation: Discount % vs Ratings")
plt.show()
          Correlation: Discount % vs Ratings
DiscountPercent
                                                                 - 0.8
                                          -0.056
                                                                 - 0.6
                                                                 - 0.4
Ratings
              -0.056
                                                                 - 0.2
                                                                 - 0.0
         DiscountPercent
                                          Ratings
```



### Myntra Exploratory Data Analysis

### Top Brands By No. of Products



SHUBHAM SHARMA



# Products Variety by Gender



SHUBHAM SHARMA

#### Pricing & Discount Strategy

### SQL Queries

Highest average discount by brand?

```
SELECT brandname , AVG(discount_percent) AS avg_discount FROM myntra
GROUP BY brandname
ORDER BY avg_discount DESC;
```

Are higher discounts linked to better ratings?

```
SELECT discountoffer , AVG(ratings) AS avg_rating , COUNT(reviews)
FROM myntra
GROUP BY discountoffer
ORDER BY discountoffer DESC;
```

	brandname character varying (300)	avg_discount numeric
1	PLATINUM Studio	99.9700000000000000
2	Lawman pg3	99.8900000000000000
3	JAIPURI BUNAAI	99.8675000000000000
4	GILLORI	99.8100000000000000
5	Teakwood Leathers	99.7983333333333333
6	ROZVEH	99.7200000000000000
7	Marvel Avengers	99.693333333333333

	discountoffer numeric	avg_rating numeric	total_reviews bigint
1	999.0	4.166666666666667	6
2	998.0	3.8000000000000000	2
3	997.0	4.3000000000000000	1
4	996.0	4.0833333333333333	6
5	995.0	4.2200000000000000	5
6	994.0	4.4000000000000000	1
7	993.0	4.066666666666667	3

#### **Pricing & Discount Strategy**

### SQL Queries

Average discount per gender (Men, Women, etc.)

```
SELECT gender , round(AVG(discountoffer),2) AS avg_discount
FROM myntra
GROUP BY gender
ORDER BY avg_discount DESC;
```

	gender character varying (20)	avg_discount numeric
1	Women	62.00
2	Men	61.98



#### **Pricing & Discount Strategy**

### SQL Queries

originalprice

numeric

avg\_price

numeric

Overpriced products compared to their average.

```
SELECT * FROM
(SELECT product_id , brandname , individual_category , originalprice,
AVG(originalprice) OVER(PARTITION BY individual_category) AS avg_price
FROM myntra) t
WHERE t.originalprice > t.avg_price
ORDER BY t.originalprice DESC;
```

product\_id [PK] integer

brandname

character varying (300)

1		16872366	Masaba	lehenga-choli	90000.0	9133.8310092961487384
2		16872354	Masaba	lehenga-choli	70000.0	9133.8310092961487384
3		10898730	MOKSHA DESIGNS	lehenga-choli	49999.0	9133.8310092961487384
4		10898720	MOKSHA DESIGNS	lehenga-choli	49999.0	9133.8310092961487384
5		16814036	Masaba	kurta-sets	49000.0	4393.9348676293622142
6		10898748	MOKSHA DESIGNS	lehenga-choli	47999.0	9133.8310092961487384
7		10898722	MOKSHA DESIGNS	lehenga-choli	47999.0	9133.8310092961487384
8		16872364	Masaba	lehenga-choli	45000.0	9133.8310092961487384
9		15271008	LADUSAA	lehenga-choli	45000.0	9133.8310092961487384
1	0	10898716	MOKSHA DESIGNS	lehenga-choli	42999.0	9133.8310092961487384

individual\_category

character varying (200)



#### **Customer Satisfaction Analysis**

Best-rated categories.

```
SELECT category , AVG(ratings) AS avg_ratings
FROM myntra
GROUP BY category
ORDER BY avg_ratings DESC
LIMIT 5;
```

Top brands by customer satisfaction

```
SELECT brandname , round(AVG(ratings),2) AS avg_rating ,
COUNT(reviews) AS total_reviews
FROM myntra
GROUP BY brandname
HAVING COUNT(*) > 20
ORDER BY avg_rating DESC , total_reviews DESC;
```

### SQL Queries

	category character varying (300)	avg_ratings numeric
1	Sports Wear	4.1505376312108075
2	Inner Wear & Sleep Wear	4.1370593073886517
3	Plus Size	4.1200574744103700
4	Lingerie & Sleep Wear	4.1096325663294007
5	Topwear	4.1059081510018086

	brandname character varying (300)	avg_rating numeric	total_reviews bigint
1	Sztori Disney	4.51	22
2	I like me	4.43	144
3	EXTRA LOVE BY LIBAS	4.37	259
4	Trident	4.37	60
5	Jockey	4.34	2224



#### **Customer Satisfaction Analysis**

## SQL Queries

Best rated sizes ?

```
SELECT sizeoption , round(AVG(ratings),2) AS avg_rating
FROM myntra
GROUP BY sizeoption
ORDER BY avg_rating DESC
LIMIT 10;
```

	sizeoption text	avg_rating numeric
1	30/32, 31/32, 32/32, 33/32, 34/32, 36/34, 36/32, 38/34	5.00
2	32.5, 34, 36, 38, 40	5.00
3	30, 32, 38, 40, 34, 36	5.00
4	28, 32, 34, 36, 40, 38, 44, 48	5.00
5	28, 38, 40	5.00
6	24, 26, 29, 32, 35, 38, 43	5.00
7	2930, 3030, 3032, 3132, 3232, 3230, 3332, 3432, 3632, 3832	5.00
8	2930, 3032, 3030, 3130, 3232, 3230, 3330, 3432, 3632	5.00
9	27, 29, 31, 33, 35, 38.5, 42, 45, 48	5.00
10	32B, 32C, 32D, 34B, 34C, 34D, 34E, 36B, 36C, 36D, 36E, 38B, 38C, 38D, 38E, 40B, 40C, 40D, 40E, 42B, 42C, 44B	5.00



#### **Customer Satisfaction Analysis**

### SQL Queries

Products that disappoint customers

```
SELECT product_id, brandname, individual_category, ratings, reviews, discount_percent
FROM myntra
WHERE ratings < 3 AND reviews > 100
ORDER BY reviews DESC;
```

	product_id [PK] integer	brandname character varying (300)	individual_category character varying (200)	ratings numeric	reviews numeric	discount_percent numeric
1	11326354	Roadster	shirts	2.9	405.0	77.49
2	2524226	HERE&NOW	jeans	2.8	382.0	80.89
3	17021492	Clovia	lounge-pants	2.6	345.0	80.81
4	11325926	Roadster	shirts	2.5	253.0	85.94
5	10187157	WROGN	jackets	2.9	239.0	94.02
6	11326204	Roadster	shirts	2.4	232.0	87.10
7	12941732	Vero Moda	tops	2.0	223.0	87.60
8	1822307	Moda Rapido	jeans	2.7	221.0	82.99
9	13841604	DOOR74	track-pants	2.9	211.0	87.58
10	11315064	MABISH by Sonal Jain	shrug	2.8	190.0	87.73



#### **Brand Performance Analysis**

### SQL Queries

Top brands by number of SKUs in each category

```
SELECT brandname , individual_category , COUNT(*) AS sku_count
FROM myntra
GROUP BY brandname , individual_category
ORDER BY individual_category , sku_count DESC ;
```

Brands relying on heavy discounts.

```
SELECT brandname ,
round(AVG(discount_percent),2) AS avg_discount ,
COUNT(*) AS total_products
FROM myntra
GROUP BY brandname
HAVING AVG(discount_percent) > 40
ORDER BY avg_discount DESC;
```



	brandname character varying (300)	individual_category character varying (200)	sku_count bigint
1	Ms.Lingies	baby-dolls	159
2	Klamotten	baby-dolls	151
3	FashionRack	baby-dolls	98
4	Clovia	baby-dolls	85
5	N-Gal	baby-dolls	64
6	XIN	baby-dolls	41
_	7.		

	brandname character varying (300)	avg_discount numeric	total_products bigint
1	PLATINUM Studio	99.97	1
2	Lawman pg3	99.89	1
3	JAIPURI BUNAAI	99.87	4
4	GILLORI	99.81	2
5	Teakwood Leathers	99.80	12
6	ROZVEH	99.72	1
7	Marvel Avengers	99.69	3
8	echt	99.67	2
9	WHITE FIRE	99.66	77
10	Sibi	99.63	3

#### **Inventory Optimization**

### SQL Queries

Slow-moving products (low demand + bad rating)

```
SELECT brandname , product_id , individual_category , ratings , reviews
FROM myntra
WHERE reviews <50 AND ratings < 3;</pre>
```

Fast-selling SKUs.

	brandname character varying (300)	product_id [PK] integer	individual_category character varying (200)	ratings numeric	reviews numeric
1	Nike	1588831	tracksuits	2.5	44.0
2	emeros	14412080	jumpsuit	2.9	33.0
3	Pepe Jeans	70172	jeans	2.9	11.0
4	KALINI	14817616	sarees	2.9	10.0
5	Clovia	9380459	bra	2.1	10.0
6	DODO & MOA	11878550	tops	2.1	8.0
7	Zink London	14600754	tops	2.4	8.0
8	abof	14473748	shirts	2.7	6.0
9	Celio	13552220	trousers	2.8	5.0
10	Huetrap	11427116	tshirts	2.8	5.0

individual\_category ratings numeric reviews numeric

<pre>SELECT brandname , product_id ,</pre>
<pre>individual_category , ratings ,</pre>
reviews , discount_percent
FROM myntra
WHERE reviews >500 OR discount_percent > 50
ORDER BY reviews DESC;

1	Mast & Harbour	2490950	tops	4.4	999.0	-66.78
2	Zivame	4335679	shapewear	4.2	999.0	22.86
3	Roadster	11895958	shirts	4.3	999.0	28.59
4	Roadster	2296012	jeans	3.9	999.0	33.36
5	LOCOMOTIVE	13780156	track-pants	4.0	999.0	13.05
6	Roadster	11690882	tshirts	4.2	999.0	-66.78
7	Mayra	8439415	tops	3.7	998.0	28.46
8	HIGHLANDER	6744434	trousers	3.9	998.0	33.42
9	Roadster	17381394	tshirts	4.3	997.0	9.20
10	Athena	11634538	jumpsuit	4.3	996.0	60.14

product\_id [PK] integer

brandname



### SQL Queries

Highest engagement categories ?

```
SELECT individual_category , SUM(reviews) AS total_reviews
FROM myntra
GROUP BY individual_category
ORDER BY total_reviews DESC
LIMIT 10;
```

	individual_category character varying (200)	total_reviews numeric
1	tshirts	4065002.335272986924432
2	sarees	3518738.814318425207032
3	kurtas	2521857.137470327448000
4	tops	2367101.351049303569968
5	dresses	2174905.268071339999184
6	shirts	2145727.967018885322528
7	jeans	1803165.629624183327104
8	kurta-sets	1659623.558872339944464
9	trousers	1490734.539797911868112
10	bra	1131950.814129361575592



### SQL Queries

Do cheaper products really get better ratings?

```
CASE
WHEN originalprice>1000 THEN 'Low Price'
WHEN originalprice BETWEEN 1000 AND 3000 THEN 'Mid Price'
ELSE 'Premium'
END AS price_segment,
round(AVG(ratings),2) AS avg_rating,
COUNT(*) AS total_products
FROM myntra
GROUP BY price_segment
ORDER BY avg_rating DESC;
```



	price_segment text	avg_rating numeric	total_products bigint
1	Premium	4.12	89113
2	Low Price	4.09	437235
3	Mid Price	4.06	214

### SQL Queries

Which brands are dominating user engagement (reviews + ratings combined)?

```
SELECT brandname,
SUM(reviews) AS total_reviews ,
round(AVG(ratings),2) AS avg_rating,
COUNT(*) AS total_products
FROM myntra
GROUP BY brandname
ORDER BY total_reviews DESC , avg_rating DESC
LIMIT 10;
```

	brandname character varying (300)	total_reviews numeric	avg_rating numeric	total_products bigint
1	Roadster	1549850.309329243951264	4.15	10935
2	Pothys	991427.201121778008184	4.09	16005
3	HERE&NOW	623192.203579606324840	4.05	6515
4	Mast & Harbour	605740.417242610758224	4.13	5148
5	KALINI	581190.656450223708632	4.06	9589
6	HRX by Hrithik Roshan	436731.352351742535408	4.16	5297
7	Sangria	401821.138058525720240	4.08	4439
8	DressBerry	398287.179820599540816	4.17	4640
9	HIGHLANDER	338105.943659013086976	4.00	2413
10	WROGN	333595.830346826879144	4.14	2471



### SQL Queries

Which categories sell at higher prices but still maintain good ratings?



	individual_category character varying (200)	avg_price numeric	avg_rating numeric	total_reviews numeric
1	lehenga-choli	9133.83	4.08	164823.793983572459440
2	coats	5637.10	4.01	23617.677436296031272
3	dress-material	5032.78	4.07	103470.917736277122416
4	burqas	4443.95	4.07	2879.596075877570656
5	kurta-sets	4393.93	4.06	1659623.558872339944464
6	jackets	4107.17	4.11	660078.786778144224632
7	co-ords	4005.30	4.07	17734.274142385983712
8	sarees	3850.75	4.08	3518738.814318425207032
9	tracksuits	3726.04	4.04	64677.904417788793536
10	nehru-jackets	3617.16	4.05	152479.721173035308616







CUSTOMER SATISFACTION







CATEGORY TRENDS Average Price

2.41K

Average Discount

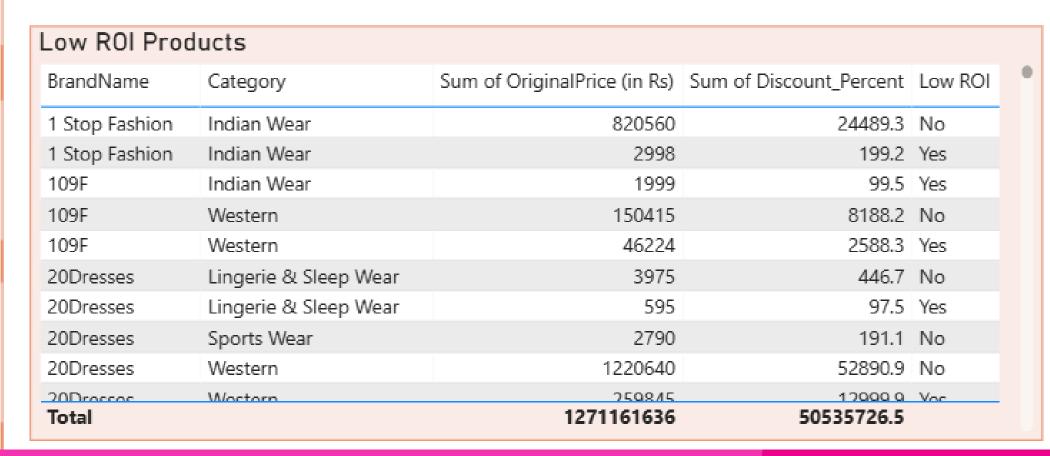
0.10K

Average Rating

4.09



Overpriced	Products		
Category	Individual_category	OriginalPrice (in Rs)	Avg_price
Bottom Wear	jeans	600	600.00
Bottom Wear	jeans	750	750.00
Bottom Wear	jeans	799	799.00
Bottom Wear	jeans	800	800.00
Bottom Wear	jeans	899	899.00
Bottom Wear	jeans	900	900.00
Bottom Wear	jeans	999	999.00
Bottom Wear	jeans	1000	1,000.00
Total			2,414.07











CUSTOMER SATISFACTION







CATEGORY TRENDS Overall Average Rating

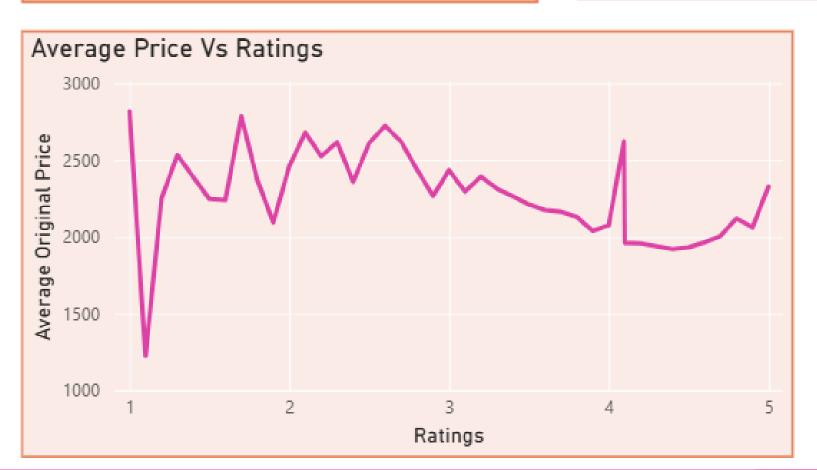
4.09

**Best Rated Category** 

### Sports Wear

Best Rated Individual Category

### thermal-tops













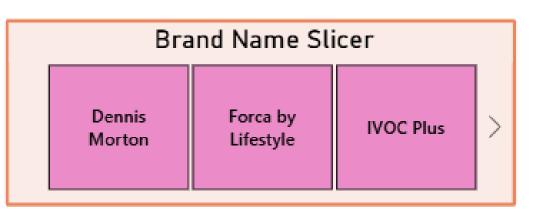
CUSTOMER SATISFACTION







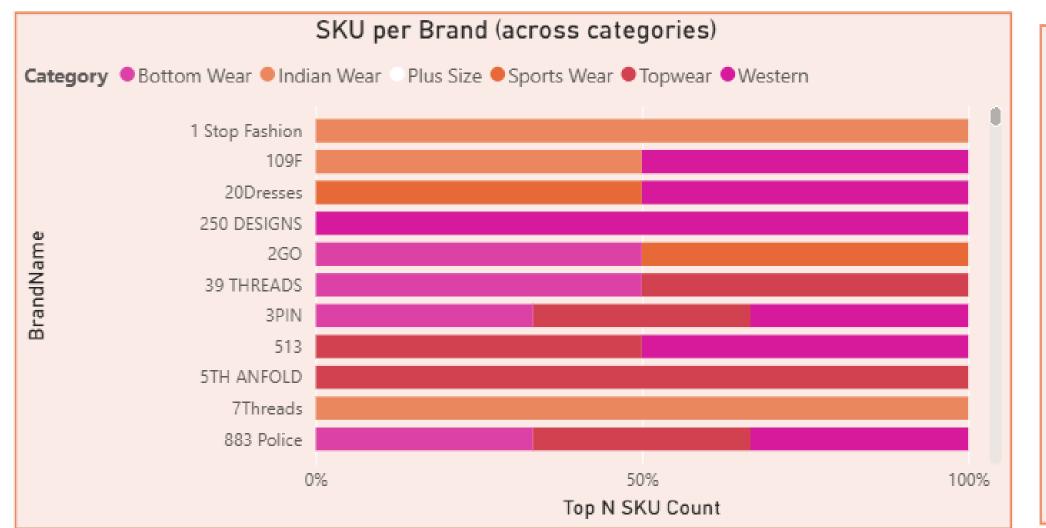
CATEGORY TRENDS



Top Brand By SKUs

Pothys













CUSTOMER SATISFACTION



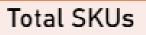
BRAND ERFORMANCE



INVENTORY OPTIMIZATION



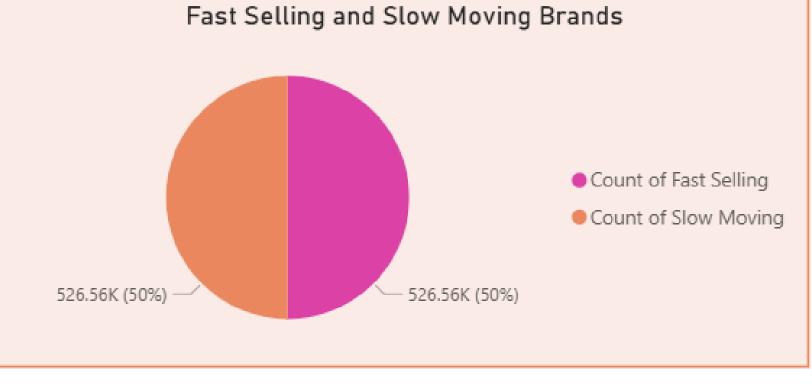
CATEGORY TRENDS

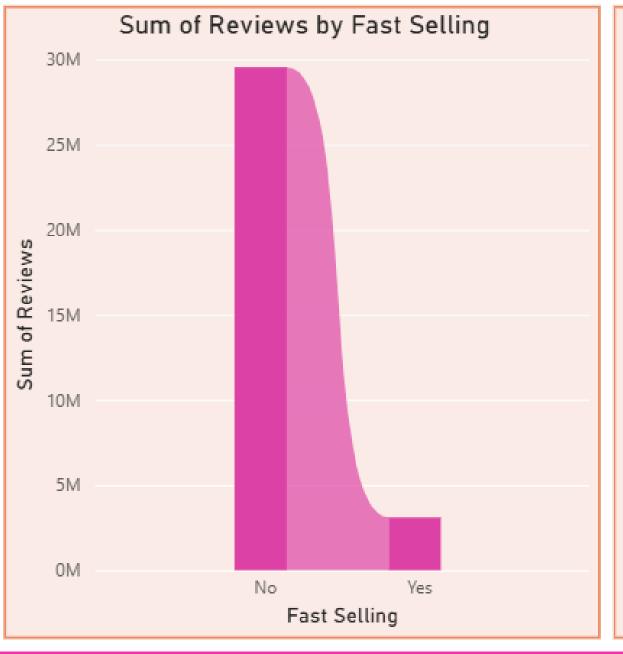


2087

Average Reviews Per SKUs

62.00





Slow Moving and Overpriced Brandname by Category				
Slow Moving	Overpriced	BrandName	Category	
No	No	Dennis Morton	Indian Wear	
No	No	Forca by Lifestyle	Bottom Wear	
No	No	IVOC Plus	Inner Wear & Sleep Wear	
No	No	IVOC Plus	Topwear	
No	No	Minions by Dressberry	Lingerie & Sleep Wear	
No	No	Newfeel By Decathlon	Inner Wear & Sleep Wear	
No	No	Powerpuff Girls by Dressberry	Lingerie & Sleep Wear	
No	No	Ratan Creation	Indian Wear	
No	No	Sztori Disney	Western	
No	No	Sztori Superman	Topwear	
No	Yes	IVOC Plus	Bottom Wear	
No	Yes	Minions by Dressberry	Lingerie & Sleep Wear	
No	Yes	PLATINUM Studio	Topwear	
No	Yes	Sztori DC	Topwear	
No	Yes	Sztori Disney	Western	
No	Yes	Sztori Superman	Topwear	
No	Yes	Zelocity	Lingerie & Sleep Wear	







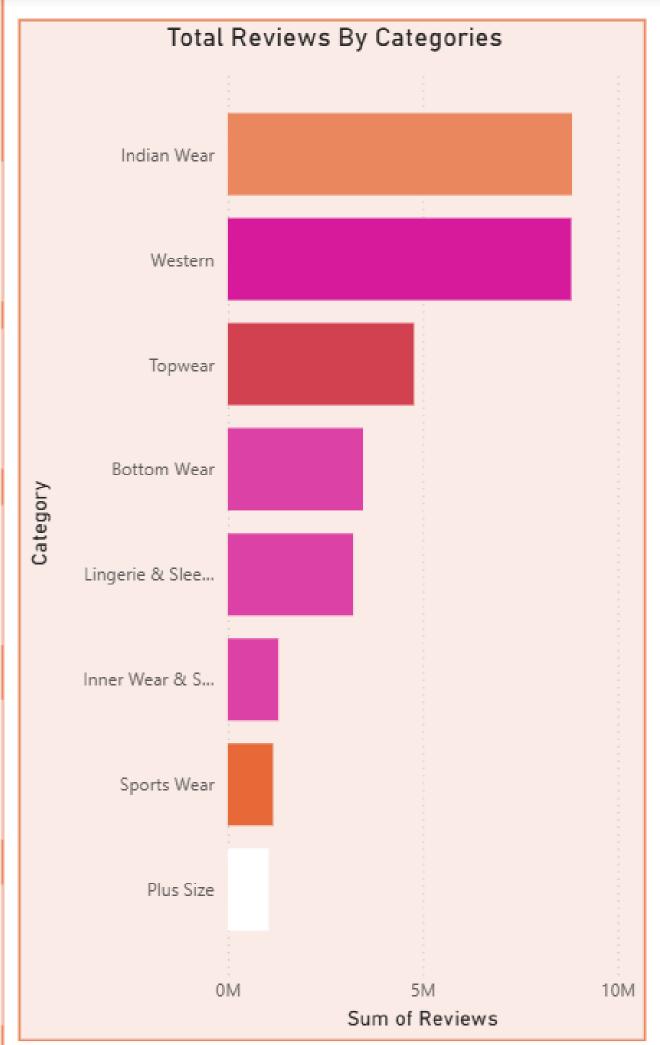
CUSTOMER SATISFACTION







CATEGORY TRENDS



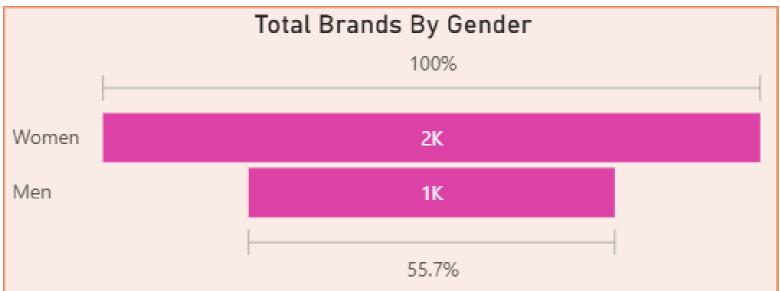
Number of Mens Product

187K

Number of Womens Product

339K





### **Analysis Report**



- Pricing & Discounts Discounts are clearly working to attract shoppers, but some brands are leaning on them a bit too much. That may win sales in the short term, but it risks eating into profitability. We also found products that are overpriced compared to their category average—suggesting Myntra needs a smarter, more balanced pricing approach.
- Customer Satisfaction The overall average rating is 4.1/5, but categories like footwear fall to 3.6, signaling quality or fit issues. Size analysis revealed that "M" sizes perform 10% better in ratings compared to "XS" or "XL," suggesting fit inconsistency.
- Brand Performance The top 5 brands contribute nearly 40% of SKUs, while smaller brands rely on >50% average discounts to drive sales. Premium brands with ratings above 4.5 manage to sell with <10% discount, proving brand trust reduces price pressure.</li>
- Inventory Optimization Roughly 18% of products are slow-moving (low reviews & low ratings), while the top 10% fast movers account for nearly 45% of customer reviews. This imbalance shows the need for sharper stock planning.
- Category Trends & Engagement Women's products make up 65% of total SKUs, but men's products are growing at ~15% year-on-year. Categories like t-shirts and dresses generate the highest engagement, with 20K+ reviews each, highlighting strong customer attention.



# THANK YOU

FOR YOUR TIME AND ATTENTION

in linkedin.com/in/shubham-shar<mark>ma190</mark>

github.com/shubhamshar19

SHUBHAM SHARMA