

Myntra Fashion Clothing EDA

Project Overview

This project performs **Exploratory Data Analysis (EDA)** on the Myntra Fashion Clothing.csv dataset to extract meaningful insights into pricing, discounts, ratings, categories, and brand distributions. The goal is to understand **customer trends and product pricing strategies** using Python.

Tools & Libraries

- **Python 3.x**
 - **Pandas** → Data manipulation and cleaning
 - **NumPy** → Numerical operations
 - **Matplotlib & Seaborn** → Data visualization
 - **4. Seaborn** → Enhanced data visualization and statistical graphics
-

Dataset Description

The dataset contains clothing product details from Myntra.

- **Brand** → Clothing brand name
 - **Category / Individual_category** → Product type (e.g., T-shirt, Saree, Dress)
 - **Gender** → Intended gender (Men/Women/Unisex)
 - **OriginalPrice** → Original listed price (in INR)
 - **DiscountedPrice** → Price after discount
 - **DiscountOffer** → Discount percentage or offer text
 - **Rating** → Customer rating
 - **Link**: <https://www.kaggle.com/datasets/manishmathias/myntra-fashion-dataset>
-

Exploratory Data Analysis (EDA) Steps

1. Data Cleaning

- Checked null values using `.info()` and `.isnull().sum()`.
- Handled missing values in `DiscountOffer`, `'DiscountPrice'` and `'Reviews'` by filling with 0 (no discount).

2. Univariate Analysis

- Histograms for price distribution.

- Boxplot shows outliers in high price range.
- Barplot for Top 10 Categories
- Countplot for Distribution by gender

3. Bivariate Analysis

- Scatterplots: Price vs Rating.
- Boxplots: Category vs Price (top 10 categories).

4. Multivariate Analysis

- Pairplot for price, discount, and rating.
- Heatmap for correlation analysis.

Key Insights

- **Category Pricing:**
 - Sarees, Kurtas, and Dresses show **higher median prices** and a wide spread → premium & budget segments both exist.
 - Casual wear like T-shirts, Jeans, and Tops show **lower and more consistent pricing**.
 - Bras are the most **affordable and uniform** category.
- **Brand Distribution:**
 - Top brands include Pothys, Roadster, and Kalini, indicating dominance in the affordable fashion segment.
- **Gender Distribution:**
 - Women's fashion has a higher product count compared to Men's.
- **Discount Trends:**
 - Discounts are higher for mid-priced products, likely to drive sales.
- **Correlation:**
 - Price and DiscountedPrice are strongly correlated.
 - Ratings have weak correlation with price, showing that affordable products can also earn high ratings.

Conclusion

The EDA reveals a **clear separation between budget and premium categories**, with ethnic/formal wear tending to be more expensive. Brands use discounts strategically, especially in the mid-range segment. These insights can help in **pricing strategies, product positioning, and marketing campaigns**.