**Flipkart E-commerce Data Analysis Using Power BI**

**AIM:**

The goal of this project is to analyze product ratings, pricing, and brand performance on Flipkart using Power BI to derive actionable insights for business decisions

**Dataset Overview**

The dataset includes product details like brand, discounted price, retail price, product ratings, and whether the product is part of Flipkart Advantage.

Source: Flipkart e-commerce data from kaggle

**Data Fields:** Brand, Product Name, Discounted Price, Retail Price, Product Rating, Flipkart Advantage.

**Methodology:**

Data Cleaning[PYTHON]: The data was cleaned to remove null values, duplicates, and outliers to ensure accuracy.

**import pandas as pd**

**# Load the dataset**

**file\_path = 'C:/Users/Administrator/Documents/bi/flipkart\_com-ecommerce\_sample.csv'**

**data = pd.read\_csv(file\_path)**

**# Drop unnecessary columns**

**data\_cleaned = data.drop(columns=['uniq\_id', 'crawl\_timestamp', 'product\_url', 'image', 'product\_specifications'])**

**# Handle missing values: replace "No rating available" with a default value of 0 for ratings**

**data\_cleaned['product\_rating'].replace('No rating available', 0, inplace=True)**

**data\_cleaned['overall\_rating'].replace('No rating available', 0, inplace=True)**

**# Convert rating columns to numeric (if applicable) and prices to floats**

**data\_cleaned['product\_rating'] = pd.to\_numeric(data\_cleaned['product\_rating'], errors='coerce').fillna(0)**

**data\_cleaned['overall\_rating'] = pd.to\_numeric(data\_cleaned['overall\_rating'], errors='coerce').fillna(0)**

**data\_cleaned['retail\_price'] = pd.to\_numeric(data\_cleaned['retail\_price'], errors='coerce').fillna(0)**

**data\_cleaned['discounted\_price'] = pd.to\_numeric(data\_cleaned['discounted\_price'], errors='coerce').fillna(0)**

**# Drop duplicates if any**

**data\_cleaned.drop\_duplicates(subset=['product\_name', 'brand', 'retail\_price'], inplace=True)**

**# Handle missing values for other columns (if needed)**

**data\_cleaned.fillna('Not available', inplace=True)**

**# Save the cleaned data to a new CSV file for Power BI**

**cleaned\_file\_path = 'C:/Users/Administrator/Documents/bi/flipkart\_cleaned\_for\_powerbi.csv'**

**data\_cleaned.to\_csv(cleaned\_file\_path, index=False)**

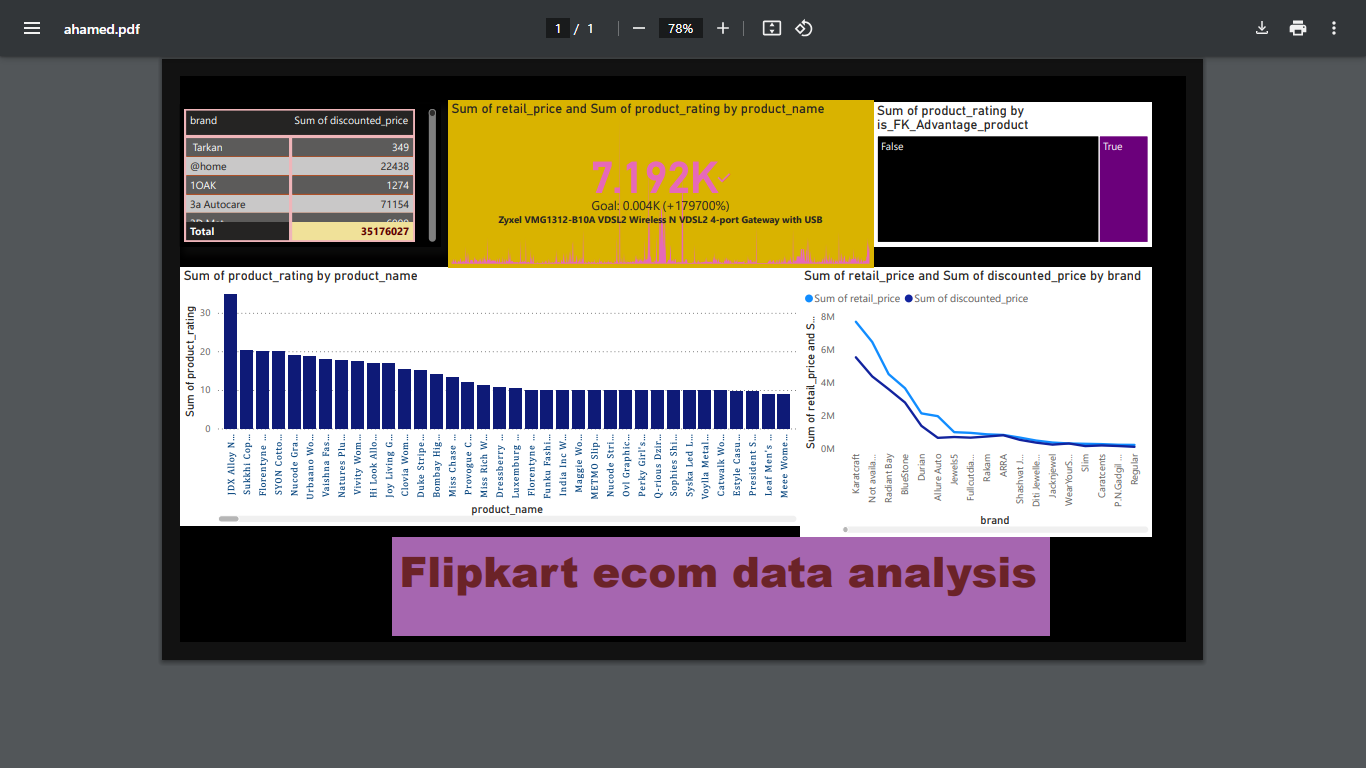
**cleaned\_file\_path**

**Data Transformation:** Calculated columns and measures were added for metrics like total sales, discount percentage, and average product rating.

**Analysis and Visualizations**

Sales by Brand: A bar chart visualizing total discounted price by brand. Key brands like @home, 3D Mat, and Tarkan contribute significantly to the total sales.

Product Rating by Name: A comparison of product ratings across different product names, indicating which products are more popular based on user feedback.



**Flipkart Advantage Products**: A pie chart or comparison showing the percentage of products under Flipkart Advantage and their impact on sales.

**Retail Price vs. Discounted Price**: A combined line or bar chart showing the comparison between the retail price and discounted price across various brands, indicating discount strategies.

**Key Findings**

Brands like @home and 3D Mat offer significant discounts, leading to higher sales.

Products under Flipkart Advantage have better ratings on average compared to others.

The highest-rated products belong to the "Florentyne" and "JD X Alloy" categories, suggesting customer satisfaction with specific brands.

**Conclusion**

The analysis shows that brands offering deep discounts tend to perform better in sales, and products with high ratings tend to belong to trusted or well-reviewed brands.