**PROJECT REPORT**



**Exploratory Data Analysis on Sales Dataset**

Submitted by:

Kashish Nigam

1. Project Objectives

- To perform exploratory data analysis (EDA) on a sales dataset.  
- To clean and preprocess the data by handling missing values and inconsistencies.  
- To visualize trends, distributions, and patterns in pricing, discounts, and savings.  
- To derive meaningful insights that can support decision-making in sales strategy.

2. Methodology and Approach

- Dataset Source: Sales dataset provided from Kaggle(Excel format).  
- Tools & Libraries Used:  
 • Python (Jupyter Notebook)  
 • Pandas, NumPy (data handling)  
 • Matplotlib, Seaborn (visualization)  
- Steps Followed:  
 1. Loaded and explored dataset structure.  
 2. Performed data cleaning (missing values, duplicates, formatting).  
 3. Created new features (Discount percentage, Savings).  
 4. Conducted univariate and bivariate analysis (distributions, correlations).  
 5. Generated visualizations (bar plots, histograms, scatter + regression).  
 6. Summarized insights and recommendations.

3. Findings and Outcomes

- Discount Trends: Higher discounts generally correlate with higher savings, but not always with higher product prices.  
- Category Insights: Certain product categories offer larger average discounts and higher savings potential.  
- Brand Insights: Top 10 brands by total savings identified — useful for promotional focus.  
- Price vs Discount: Regression analysis revealed a moderate relationship between product price and discount %.  
- Top Products:  
 • Highest discount % products were often not the highest in savings.  
 • Best value (savings) came from moderately priced items with higher discount %.

4. Learnings

In this project, I have learned various things –

* How the data is cleaned?
* How the analysis is done?
* Different features of python libraries
* How the visualizations are created?
* How to identify trends, patterns for future growth?

5. Conclusion

This EDA project provided key insights into sales, discounting strategies, and customer savings. The analysis highlighted the importance of targeted discounts by category and brand. These insights can support pricing optimization, marketing campaigns, and customer engagement strategies.

Thankyou!!