INTERNSHIP PROJECT REPORT

TITLE   - AMAZON SALES ANALYSIS

NAME - SOUNDHAR G

BATCH - 01

TABLE OF CONTENTS

INTRODUCTION:

Amazon.com, Inc. is an American multinational technology company that focuses on e-commerce, cloud computing, digital streaming, and artificial intelligence.

It is one of the Big Five companies in the U.S. information technology industry, along with Google, Apple, Microsoft, and Facebook.

The company has been referred to as "one of the most influential economic and cultural forces in the world," as well as the world's most valuable brand.

DETAILS OF DATA:

- EDA PROJECTS WE WILL BE PERFORMING AMAZON SALES DATA ANALYSIS AND IT IS A INTERESTING DATASET TO WORK ON.

- EXPLORE THE POPULARITY OF PRODUCTS BASED ON THE NUMBEROF USER REVIEWS AND RATINGS

TOOLS USED:

- PANDAS

- MATPLOTLIB

- POWER BI

MAIN KPI'S USED:

TASKS:

- IDENTIFY TOP 5 PRODUCTS BY RATINGS

- ANALYSIS THE DISTRIBUTION OF DISCOUNT ANALYSIS BASED ON RATINGS.

- IDENTIFY AVERAGING RATINGS FOR PRODUCTS IN MAIN CATEGORY.

- IDENTIFY CATEGORY WISE SALES

- IDENTIFY MOST DISCUSSED PRODUCTS BASED ON RATINGS

SCREENSHOTS:

12 3. Discount Percentage Analysis:

Task(3): Examine the distribution of discount percentages across different products. Columns:

```python
plt.figure(figsize=(12, 6))
sns.histplot(df['discount_percentage'], bins=20, kde=True, color='blue')

plt.title('Distribution of Discount Percentages')
plt.xlabel('Discount Percentage')
plt.ylabel('Frequency')
plt.show()
```

SCREENSHOTS:

14 5. Discounted Price vs. Actual Price Comparison

Task(10): Explore the distribution and relevance of product image links.

```python
df[['discounted_price', 'actual_price']] = df[['discounted_price',
 ↪'actual_price']].astype(float)


# Grouped bar chart for Discounted Price vs. Actual Price
plt.figure(figsize=(12, 6))
df[['discounted_price', 'actual_price']].mean().plot(kind='bar',
 ↪color=['skyblue', 'orange'])
```
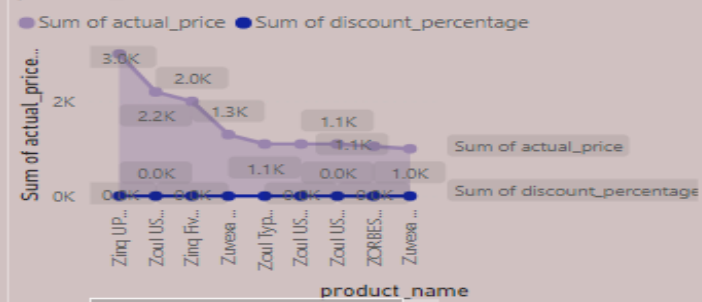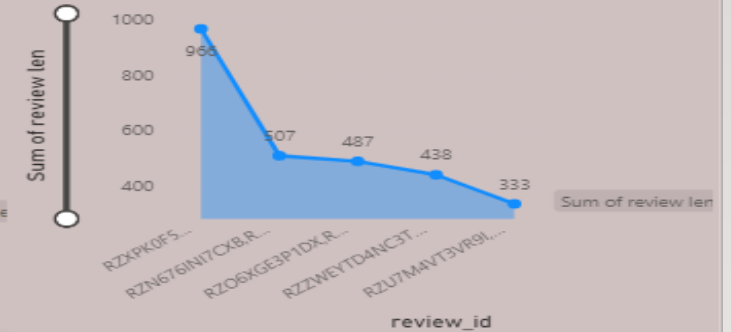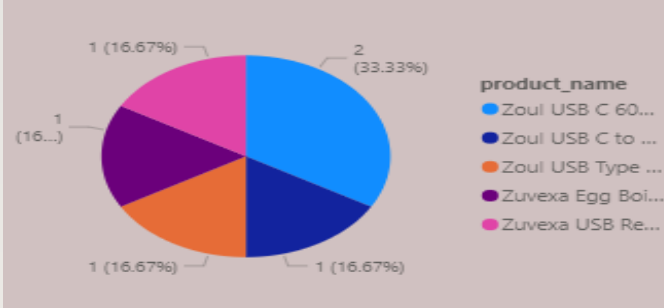
MY DESIGN:

AMAZON SALES ANALYSIS
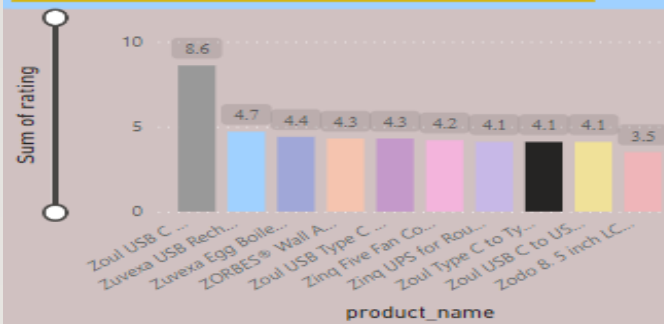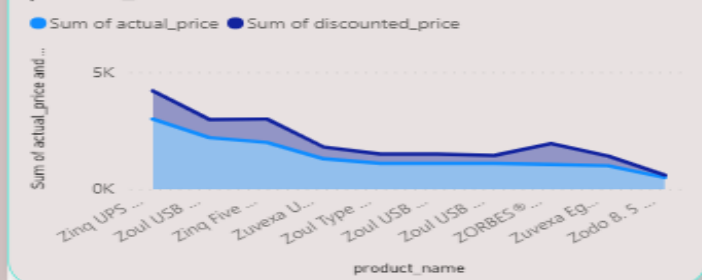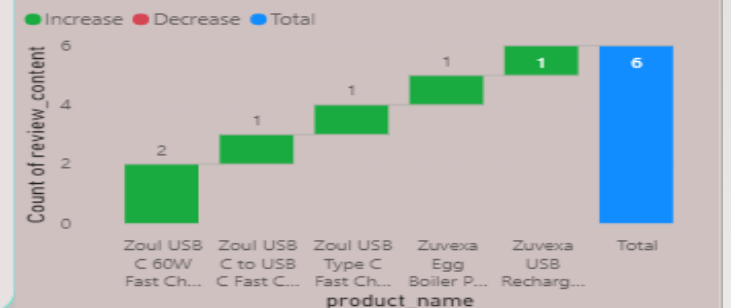
CONCLUSION:

- In conclusion, the methods employed in the data analysis of the given dataset were effective in providing insights into the data distribution.

- The challenges encountered during the implementation of the methods provided valuable lessons for future works.

THANKYOU