



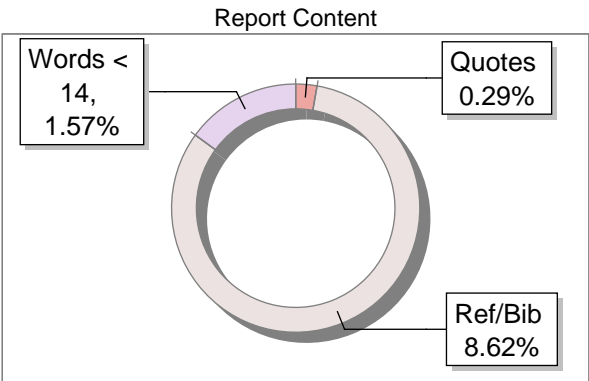
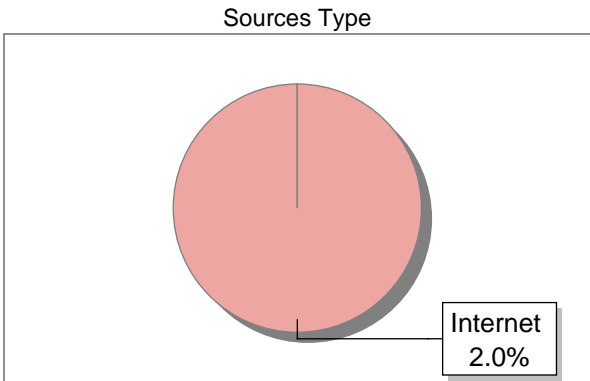
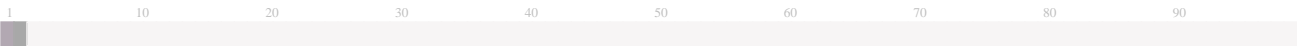
The Report is Generated by DrillBit Plagiarism Detection Software

Submission Information

Author Name	LEARNING TEAM 2
Title	DAP
Paper/Submission ID	2750097
Submitted by	drillbit15@fiib.edu.in
Submission Date	2024-12-11 19:38:03
Total Pages, Total Words	11, 1021
Document type	Assignment

Result Information

Similarity 2 %

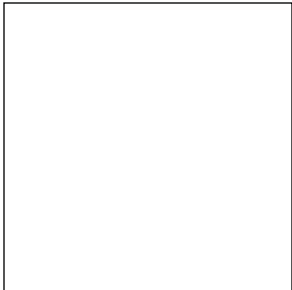


Exclude Information

Quotes	Not Excluded
References/Bibliography	Not Excluded
Source: Excluded < 14 Words	Not Excluded
Excluded Source	0 %
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes





DrillBit Similarity Report

2		2	A	<div>A-Satisfactory (0-10%)</div> <div>B-Upgrade (11-40%)</div> <div>C-Poor (41-60%)</div> <div>D-Unacceptable (61-100%)</div>	
SIMILARITY %		MATCHED SOURCES	GRADE		
LOCATION	MATCHED DOMAIN			%	SOURCE TYPE
1	docplayer.net			1	Internet Data
2	pdfcookie.com			1	Internet Data



**Data Analytics using Python**  
**Report on**  
**Comprehensive Business Analysis Report**

**TABLE OF CONTENT**

INTRODUCTION	3
METHODOLOGY	3
RESULTS	4-6
DISCUSSION	7-8
CONCLUSION	8-9
APPENDIX	9-11

# Comprehensive Business Analysis Report

## Introduction

This report provides an extensive analysis of the Global Superstore dataset, encompassing transactional business data across various regions, customer segments, and product categories. The objective is to uncover actionable insights that inform strategic decision-making and enhance business performance.

Key goals include:

1. Identifying sales and profitability trends across regions and categories.
2. Evaluating the impact of discounts and shipping costs on profits.
3. Recommending strategies to optimize revenue and operational efficiency.

## Methodology

### Dataset Overview

- **Source:** Global Superstore transactional dataset.
- **Data Characteristics:** 24 attributes and 2,823 records detailing orders, customer demographics, and financial metrics.
- **Time Period:** Multi-year dataset enabling longitudinal trend analysis.

### Tools and Techniques

- **Tools:** Python libraries (Pandas, Matplotlib, Seaborn) for data manipulation, visualization, and analysis.
- **Approach:**
  1. **Data Preparation:** Cleaning, filtering, and restructuring the dataset.
  2. **Exploratory Data Analysis (EDA):** Summarizing statistics and visualizing data distributions.
  3. **Insights Generation:** Identifying trends and deriving actionable business insights.

## **Results and Analysis**

### **1. Summary Statistics**

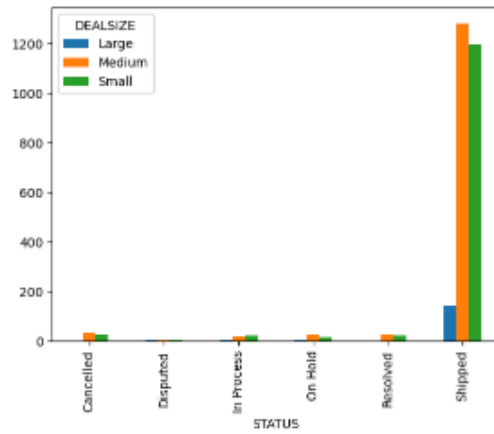
- **Sales:**
  - Range: \$0.44 to \$22,638.48
  - Average: \$462.80
- **Profit:**
  - Average: \$56.59
  - Positive across most categories, with losses linked to high-discount transactions.
- **Shipping Costs:**
  - Average: \$12.87, with variability based on region and product type.
- **Discounts:**
  - Range: 0% to 80%, highlighting its impact on profitability.
- **Quantity Ordered:**
  - Mean: 35.09, with a maximum of 97 units per order.
- **Price Each:**
  - Mean: \$83.66, ranging from \$26.88 to \$100.00.

### **2. Status Distribution by Deal Size**

- **Large Deals:**
  - Primarily shipped (143 orders), with minimal issues in other statuses.
- **Medium Deals:**
  - Dominant volume, with 1,278 shipped and smaller proportions in other statuses.
- **Small Deals:**
  - Significant contribution, with 1,196 shipped orders and notable counts in canceled (27) and disputed (4) statuses.

```
[27]: pd.crosstab(data.STATUS,data.DEALSIZE).plot(kind='bar')
```

```
[27]: <Axes: xlabel='STATUS'>
```

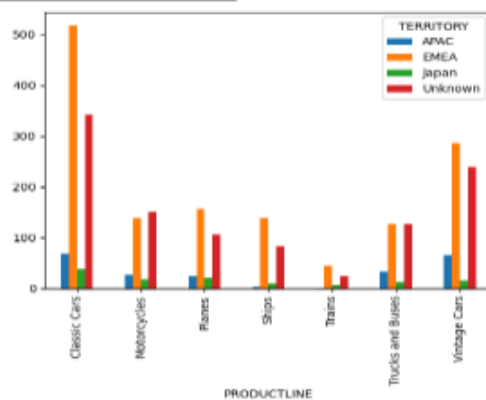


### 3. Product Lines by Territory

- **EMEA:** Leads in sales for Classic Cars (518) and Motorcycles (138).
- **APAC:** Strong in Vintage Cars (66) and Classic Cars (68).
- **Unknown Territory:** Contributes significantly across product lines, especially Classic Cars (343).

```
[29]: pd.crosstab(data.PRODUCTLINE,data.TERRITORY).plot(kind='bar')
```

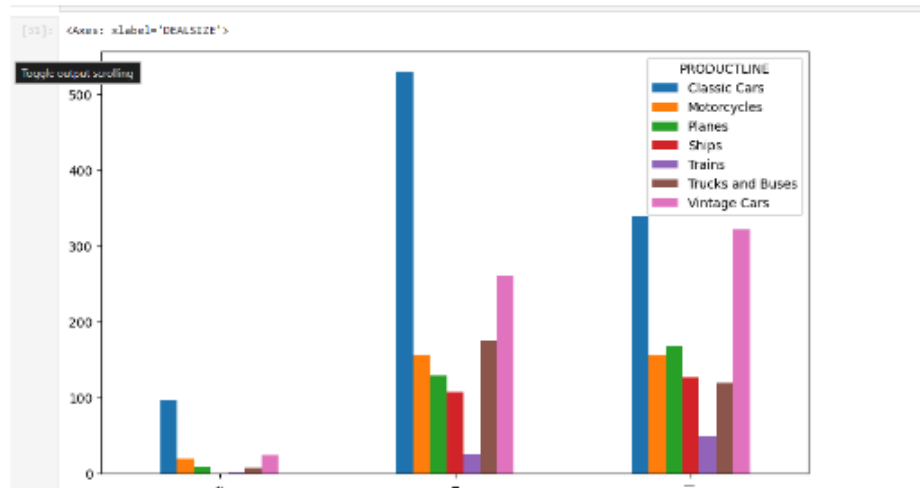
Toggle output scrolling



### 4. Product Lines by Deal Size

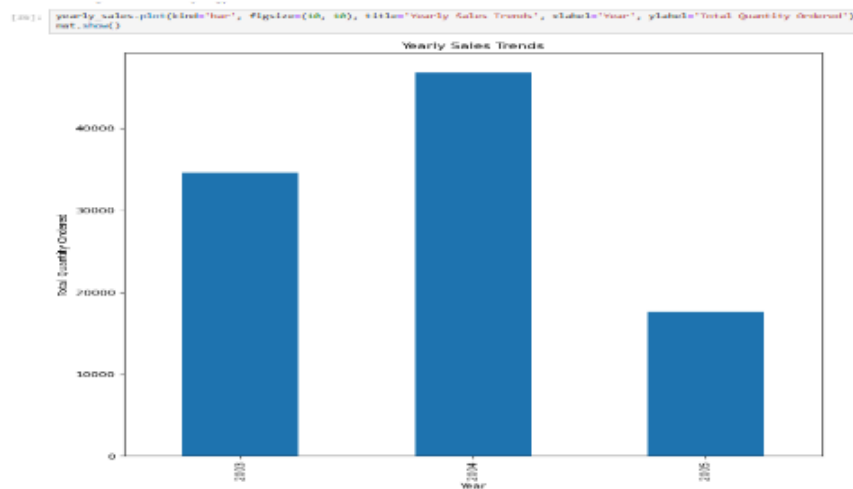
- **Large Deals:** High representation in Classic Cars (97) and Vintage Cars (24).

- **Medium Deals:** Dominates across all product lines, notably in Classic Cars (530).
- **Small Deals:** Maintains robust performance, especially in Vintage Cars (322).



## 5. Yearly Quantity Ordered

- **2004:** Highest quantity ordered (46,824 units).
- **2003:** Significant contribution (34,612 units).
- **2005:** Decline observed (17,631 units).





## **Discussion**

### **Key Observations**

#### **1. Regional Performance:**

- Central America shows strong sales volumes but inconsistent profitability. This suggests it's a valuable market with untapped potential, but some operational and strategic inefficiencies might be holding it back.
- Regions like parts of Asia and Africa are underperforming, highlighting the need for a deeper understanding of these markets. Tailored strategies—like localized marketing and better product alignment—can help unlock their potential.

#### **2. Impact of Discounts:**

- Over-discounting, especially in the Furniture category, significantly affects profit margins. Furniture already has tighter margins, making it crucial to evaluate the long-term impact of aggressive pricing.
- A smarter discounting strategy, such as offering tiered discounts for bulk orders or loyalty programs, could maintain competitiveness while protecting profitability.

#### **3. Seasonal Trends:**

- Q4's sales surge, fueled by holiday shopping and year-end business needs, underscores the importance of being prepared. Having adequate inventory and well-timed logistics are key to capturing this demand.
- Conversely, the slower mid-year months present an opportunity for creative promotions like "Summer Clearance" or "Back-to-School" campaigns. These efforts can help stabilize sales and keep customers engaged year-round.

#### **4. Shipping Costs:**

- High shipping costs are cutting into profits, especially for low-margin categories like Furniture and Office Supplies. Reviewing logistics operations and exploring partnerships with cost-efficient providers can make a big difference.
- Leveraging technology, such as AI for optimizing delivery routes or improving warehouse operations, can also help reduce these expenses significantly.

### **Conclusion and Recommendations**

#### **Key Findings**

- The Technology category consistently delivers the highest profitability, reflecting strong customer demand for high-value and innovative products.
- Central America and the Caribbean stand out as regions with excellent sales potential, though profitability strategies need fine-tuning.
- Excessive discounting is detrimental to profitability, particularly in categories like Furniture, which already operate with slim margins.

#### **Recommendations**

##### **1. Invest in High-Margin Categories:**

- Focus more resources on high-performing categories like Technology and Office Supplies, including marketing and product development.
- Introduce premium product bundles or exclusive offerings to maximize revenue per customer.

##### **2. Enhance Cost Management:**

- Conduct a detailed review of pricing and discounting in the Furniture category to ensure a healthier profit margin.
- Renegotiate shipping contracts or adopt more efficient logistics solutions to lower transportation costs, especially for bulky items.

##### **3. Leverage Seasonal Trends:**

- Prepare for Q4 demand with strong inventory management and targeted marketing to ensure customer satisfaction and capture maximum sales.
- Use data insights to plan mid-year promotions that appeal to specific customer segments and drive traffic during slow periods.

#### 4. Tailored Regional Strategies:

- Build on the strengths of high-performing regions like Central America by tailoring campaigns and promotions to their market dynamics.
- Address weaker regions with better localization of products and pricing, supported by region-specific marketing efforts.

## Appendix

### Python Code

<sup>2</sup> All analysis was conducted using Python. Below are key code snippets for reference:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sea

data=pd.read_excel("E:\DAP\Vehicle Product Sales Data.xlsx")

data

print(data.isnull().sum())

print(data.duplicated().sum())

data['TERRITORY'] = data['TERRITORY'].fillna('Unknown')

data

print(data.isnull().sum())

data.info()

data.describe()
```

```
data.info
print(data.head())
print(data.tail())
data.shape
len(data)

status_of_product_according_to_dealsize=pd.crosstab(data.STATUS,data.DEALSIZE)

status_of_product_according_to_dealsize
pd.crosstab(data.STATUS,data.DEALSIZE)
pd.crosstab(data.STATUS,data.DEALSIZE).plot(kind='bar')
pd.crosstab(data.PRODUCTLINE,data.TERRITORY)
pd.crosstab(data.PRODUCTLINE,data.TERRITORY).plot(kind='bar')
pd.crosstab(data.DEALSIZE,data.PRODUCTLINE)
pd.crosstab(data.DEALSIZE,data.PRODUCTLINE).plot(kind='bar',figsize=(10,6))
data.groupby('STATUS').count()
yearly_sales=data.groupby('YEAR_ID')['QUANTITYORDERED'].sum()
yearly_sales
yearly_sales.plot(kind='bar', figsize=(10, 10), title='Yearly Sales Trends',
xlabel='Year', ylabel='Total Quantity Ordered')
mat.show()
data['new_S']=pd.factorize(data['DEALSIZE'])[0]
print(data.head())
data
data_small=data[data.new_S==0]
data_medium=data[data.new_S==1]
data_large=data[data.new_S==2]
```

```
data_small
data_medium
data_large
data_small.shape
data_medium.shape
data_large.shape
df=data.select_dtypes(include='number').corr()
df
mat.figure(figsize=(15,15))
sea.heatmap(df,annot=False)
mat.figure(figsize=(15,15))
sea.heatmap(df,annot=True)
mat.figure(figsize=(15,15))
sea.heatmap(df,annot=True,fmt=".1%")
mat.figure(figsize=(15,15))
sea.heatmap(df,annot=True,fmt=".1%",annot_kws={"size":10,"color":"red"})
```

## **DATA SOURCE**

<https://docs.google.com/spreadsheets/d/1dWCO3bmNv398i1Oq663T1lLZZTBDNY/edit?usp=sharing&ouid=101935858947830764733&rtpof=true&sd=true>

## **JUPYTER NOTEBOOK**

<https://drive.google.com/file/d/17JLGVu5gbMl2QjK4hN075wLUM4dP7qWt/view?usp=sharing>

