Superstore Sales Analysis - Full Project Report

This report presents a full Python-based analysis of Superstore Sales Data, including data loading, cleaning, visualization, and insights. The code sections are directly included for reference. Dataset Info: - Total Records: 9994 - Total Columns: 13 - Total Sales: ■2,297,200.86 - Total Profit: ■286,397.02 - Total Quantity Sold: 37,873

Python Project Code

```
# ■ Import Libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

# ■ Load Dataset
file_path = "1719219914-Analysis of Super Store - DA (1).csv"
df = pd.read_csv(file_path)
df.head()

# ■ Data Overview
df.info()
df.isnull().sum()
df.describe()

# ■ Basic Metrics
print("Total Sales:", df['Sales'].sum())
print("Total Profit:", df['Profit'].sum())
print("Total Quantity:", df['Quantity'].sum())

# ■ Visualizations

# 1■■ Sales by Category
sns.barplot(data=df, x='Category', y='Sales', estimator=np.sum, palette='Set2')

# 2■■ Profit by Sub-Category
sns.barplot(data=df, x='Sub-Category', y='Profit', estimator=np.sum, palette='coolwarm')

# 3■■ Sales by Region
sns.barplot(data=df, x='Region', y='Sales', estimator=np.sum, palette='pastel')

# 4■■ Ship Mode Preference
df['Ship Mode'].value_counts().plot(kind='pie', autopct='%1.1f%%')

# 5■■ Sales vs Profit Relationship
sns.scatterplot(data=df, x='Sales', y='Profit', hue='Category', alpha=0.7)

# 6■■ Time Series (if 'Order Date' exists)
df['Order Date'] = pd.to_datetime(df['Order Date'])
df['YearMonth'] = df['Order Date'].dt.to_period('M')
monthly_sales = df.groupby('YearMonth')['Sales'].sum()
monthly_sales.plot(marker='o')

# ■ Insights
print("Highest Sales Category:", df.groupby('Category')['Sales'].sum().idxmax())
print("Most Profitable Sub-Category:", df.groupby('Sub-Category')['Profit'].sum().idxmax())
print("Region with Most Sales:", df.groupby('Region')['Sales'].sum().idxmax())
print("Most Used Ship Mode:", df['Ship Mode'].value_counts().idxmax())
```

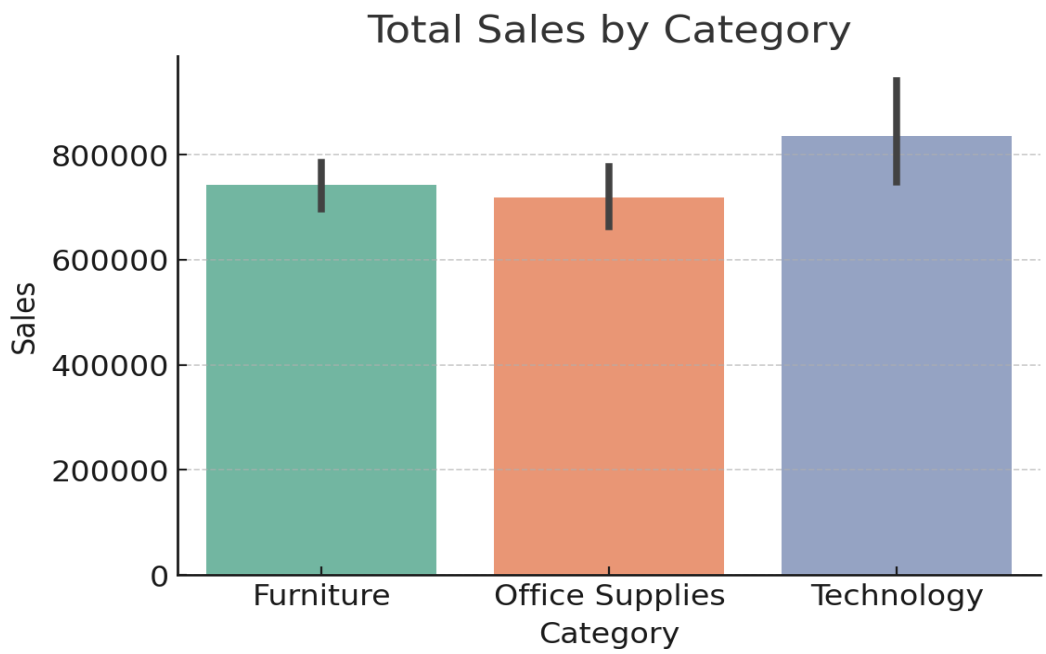

Figure 1: Total Sales by Category

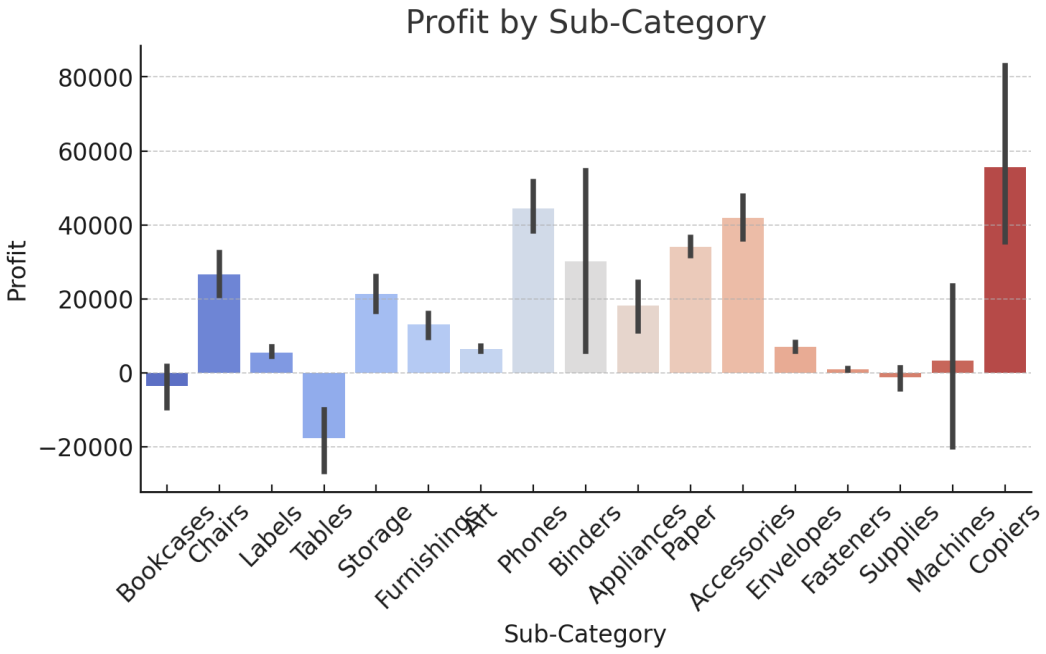

Figure 2: Profit by Sub-Category

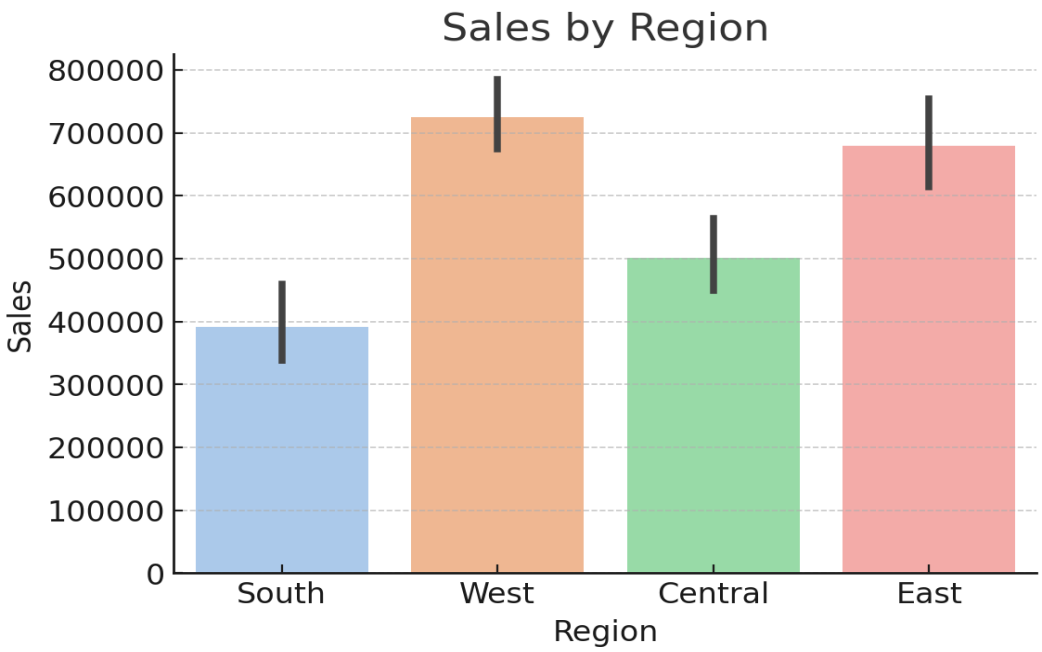

Figure 3: Sales by Region

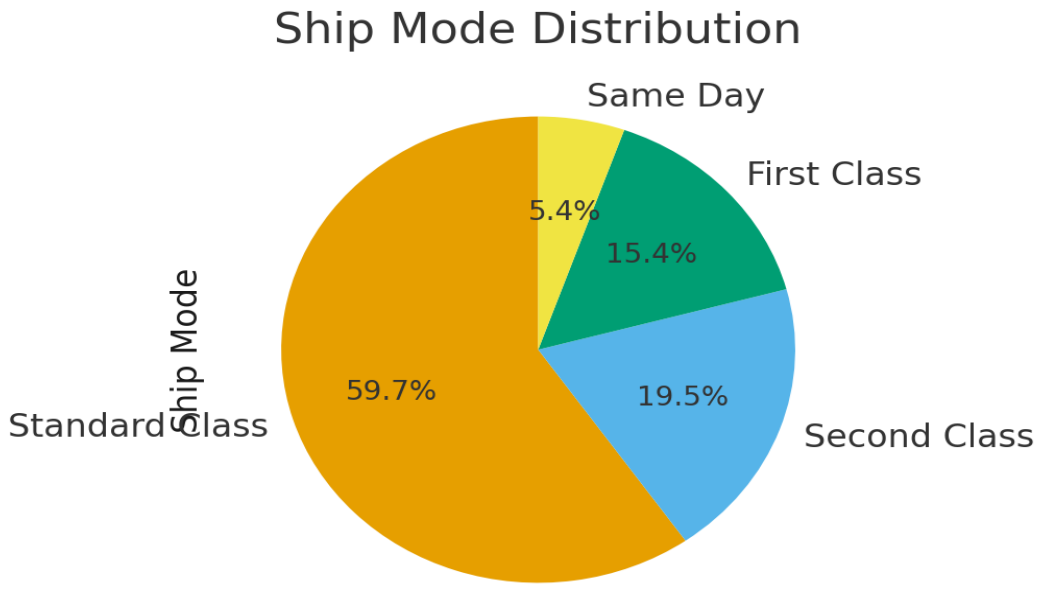

Figure 4: Ship Mode Preference

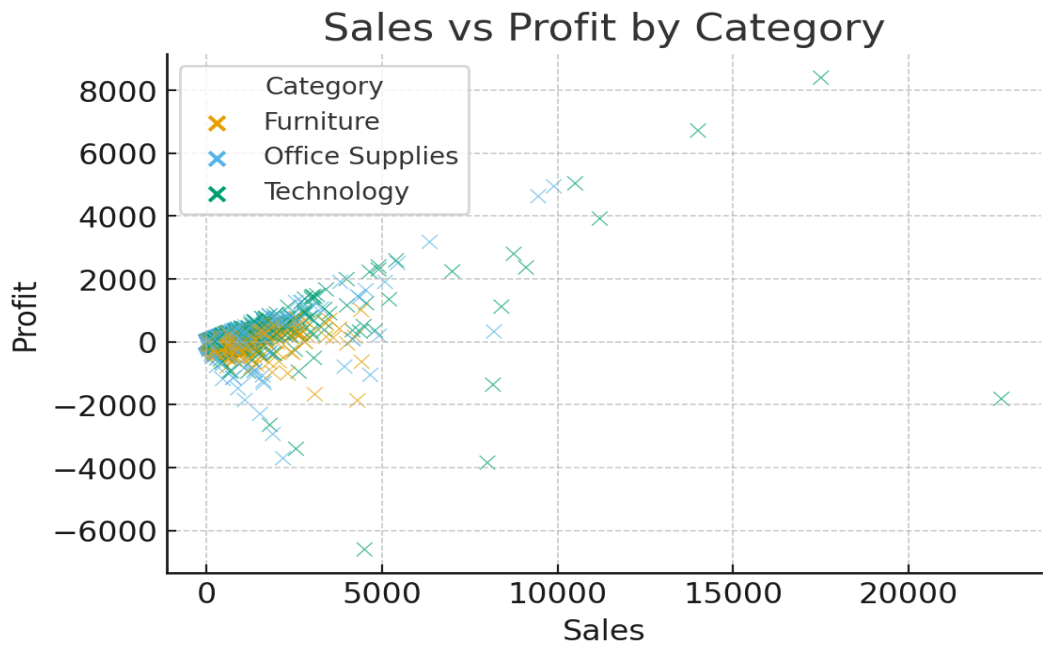

Figure 5: Relationship between Sales and Profit

Key Insights: - Highest Sales Category: Technology
- Most Profitable Sub-Category: Copiers
- Region with Highest Sales: West
- Most Used Ship Mode: Standard Class