

**Turning Raw Data into
Business Insights**

01

DATA COLLECTION REPORT

02

AVERAGE PROFIT VALUE BY DISCOUNT
BAND

03

PROFIT AMOUNT BY CATEGORY AND SUB-
CATEGORY

04

SALES TOTAL AND PROFIT TOTAL BY
REGION AND STATE

05

AMOUNT QUANTITY AND AMOUNT
PROFIT BY SUB-CATEGORY

06

TOP 10 PRODUCTS, BOTTOM 10
PRODUCTS

07

STRATEGIC RECOMMENDATIONS

08

THANK YOU

Data Collection Report

```
import pandas as pd

# Load original Excel file
df_original = pd.read_excel("Uncleaned Sample Superstore.xlsx", sheet_name="Sample Superstore")

# Make a copy for cleaning
df_cleaned = df_original.copy()

# 1. Remove duplicates
df_cleaned.drop_duplicates(inplace=True)

# 2. Fill missing values in 'Discount' with 0.0
df_cleaned['Discount'].fillna(0.0, inplace=True)

# 3. Check missing values BEFORE deletion
missing_report = df_cleaned.isnull().sum().reset_index()
missing_report.columns = ['Column', 'Missing Values']
print("Missing values before deletion:\n")
print(missing_report)

# 4. Drop rows with any remaining missing values
df_cleaned.dropna(inplace=True)

# 5. Convert Postal Code to string
df_cleaned['Postal Code'] = df_cleaned['Postal Code'].astype(str)

# 6. Reset index
df_cleaned.reset_index(drop=True, inplace=True)

# 7. Save cleaned data to a new Excel file
df_cleaned.to_excel("Cleaned_Sample_Superstore.xlsx", index=False)

# 8. Final summary
print("\nCleaned file saved as Cleaned_Sample_Superstore.xlsx")
print("Final shape:", df_cleaned.shape)
print("Missing values after cleaning:\n")
```

Missing values before deletion:

	Column	Missing Values
0	Ship Mode	0
1	Segment	1
2	Country	2
3	City	10
4	State	13
5	Postal Code	34
6	Region	1
7	Category	25
8	Sub-Category	23
9	Sales	15
10	Quantity	4
11	Discount	0
12	Profit	17

Cleaned file saved as Cleaned_Sample_Superstore.xlsx

Final shape: (9839, 13)

Missing values after cleaning:

Ship Mode	0
Segment	0
Country	0
City	0
State	0
Postal Code	0
Region	0
Category	0
Sub-Category	0
Sales	0
Quantity	0
Discount	0
Profit	0
dtype:	int64

 **SOURCE: KAGGLE - A WIDELY USED DATASET FOR BUSINESS ANALYSIS AND POWER BI PRACTICE. WE CHOSE IT BECAUSE IT REFLECTS REAL-WORLD RETAIL OPERATIONS AND INCLUDES PRODUCT, CUSTOMER, AND FINANCIAL DATA.**

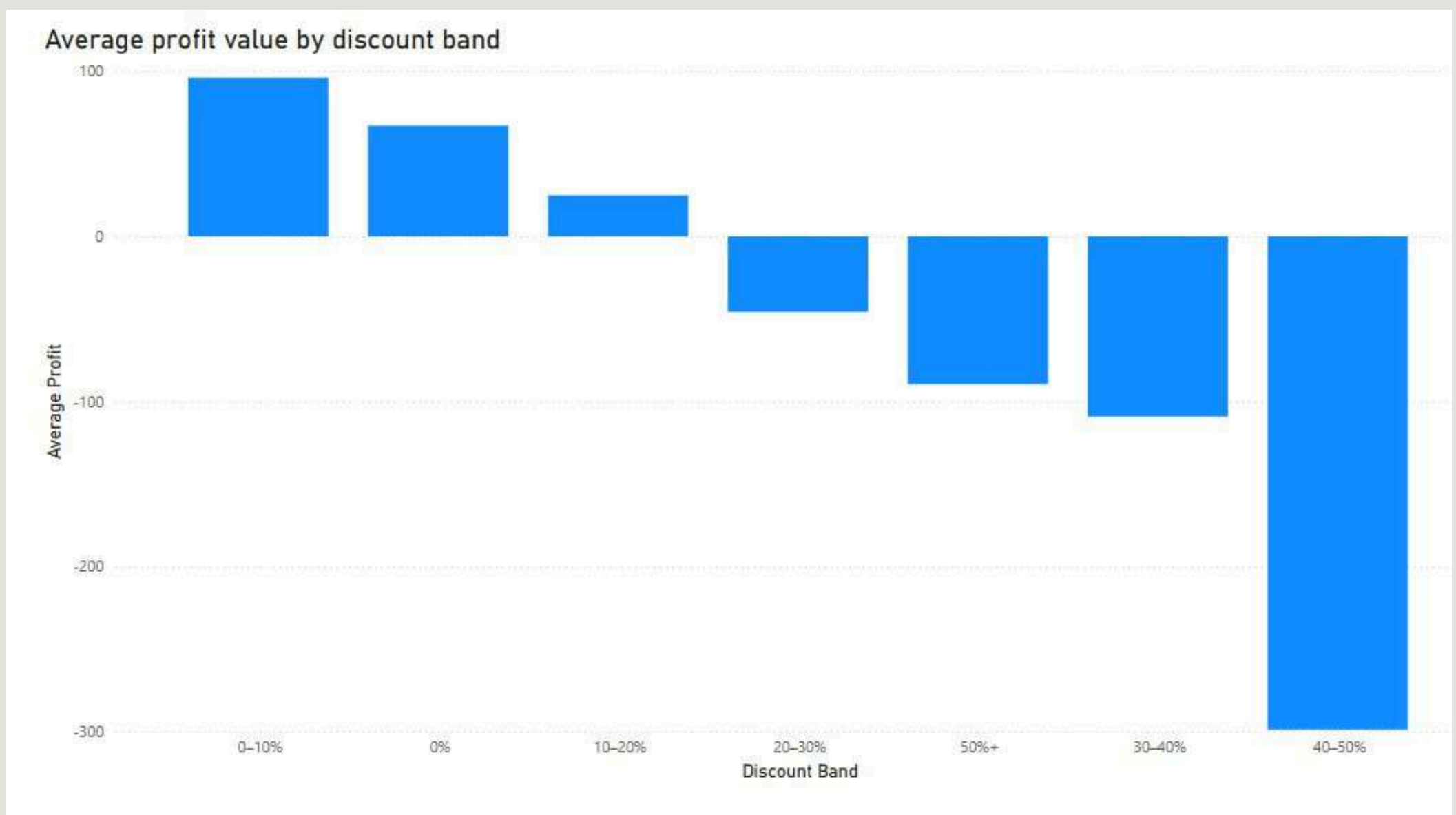
 **OVERVIEW:**

- 9,995 ROWS × 13 COLUMNS
- COVERS SALES TRANSACTIONS ACROSS U.S. STATES
- KEY FIELDS: SHIP MODE, SEGMENT, CATEGORY, SALES, QUANTITY, DISCOUNT, PROFIT

 **CLEANING STEPS (PYTHON - PANDAS):**

- REMOVED DUPLICATES
- FILLED MISSING DISCOUNTS WITH 0.0
- DROPPED ROWS WITH CRITICAL NULLS (CATEGORY, SUB-CATEGORY, QUANTITY, PROFIT)
- CONVERTED POSTAL CODE TO STRING FOR POWER BI COMPATIBILITY

Average profit value by discount band

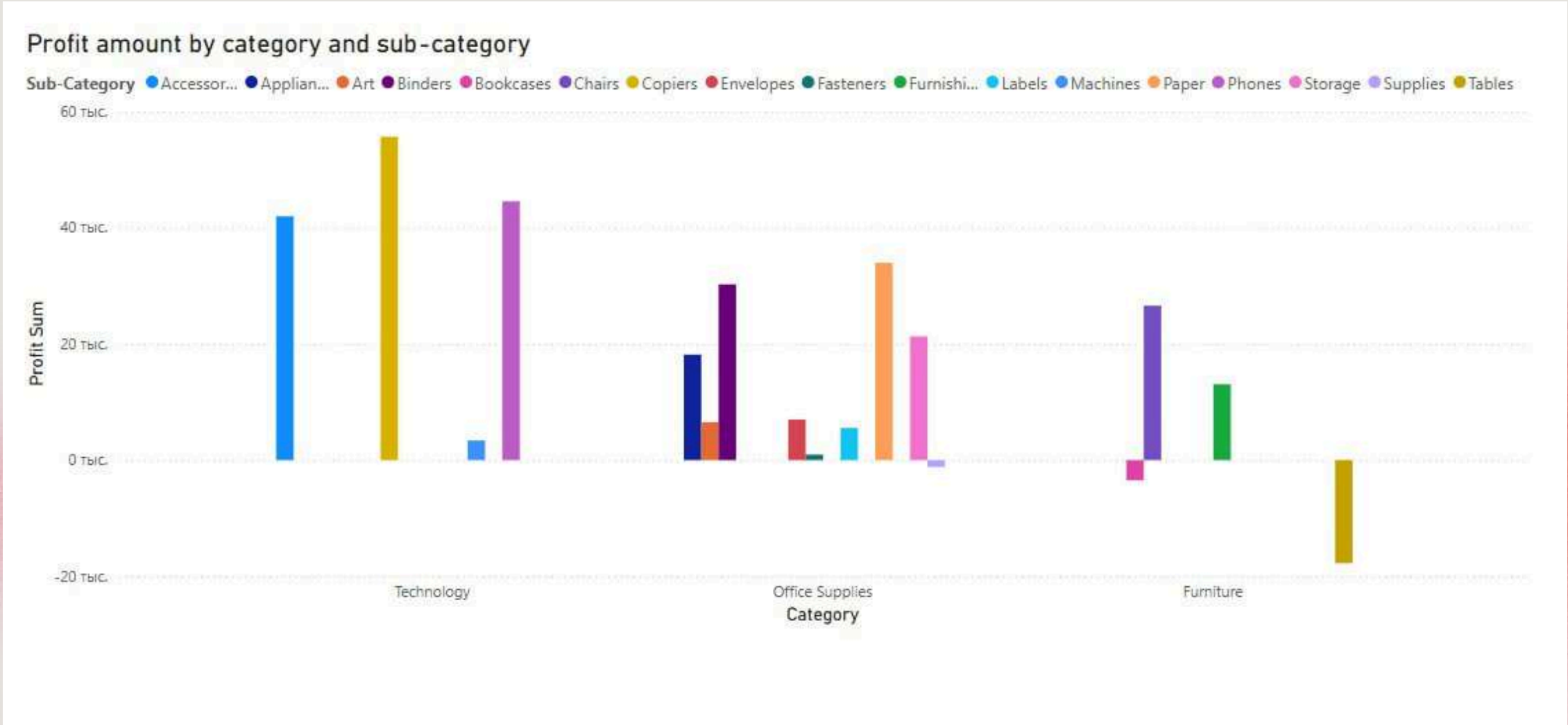


AS THE DISCOUNT PERCENTAGE INCREASES, THE AVERAGE PROFIT DECREASES SIGNIFICANTLY, INDICATING THAT HIGHER DISCOUNT BANDS ARE ASSOCIATED WITH NEGATIVE PROFITABILITY.

THE CHART SHOWS THAT THE AVERAGE PROFIT IS POSITIVE FOR LOWER DISCOUNT BANDS (0%, 0-10%, 10-20%) BUT TURNS NEGATIVE AS DISCOUNTS INCREASE (ESPECIALLY FROM 20% AND ABOVE).

THE 40-50% DISCOUNT BAND SHOWS THE MOST NEGATIVE AVERAGE PROFIT, SUGGESTING STEEP DISCOUNTS MAY BE CAUSING LOSSES RATHER THAN BOOSTING SALES VOLUME ENOUGH TO COMPENSATE.

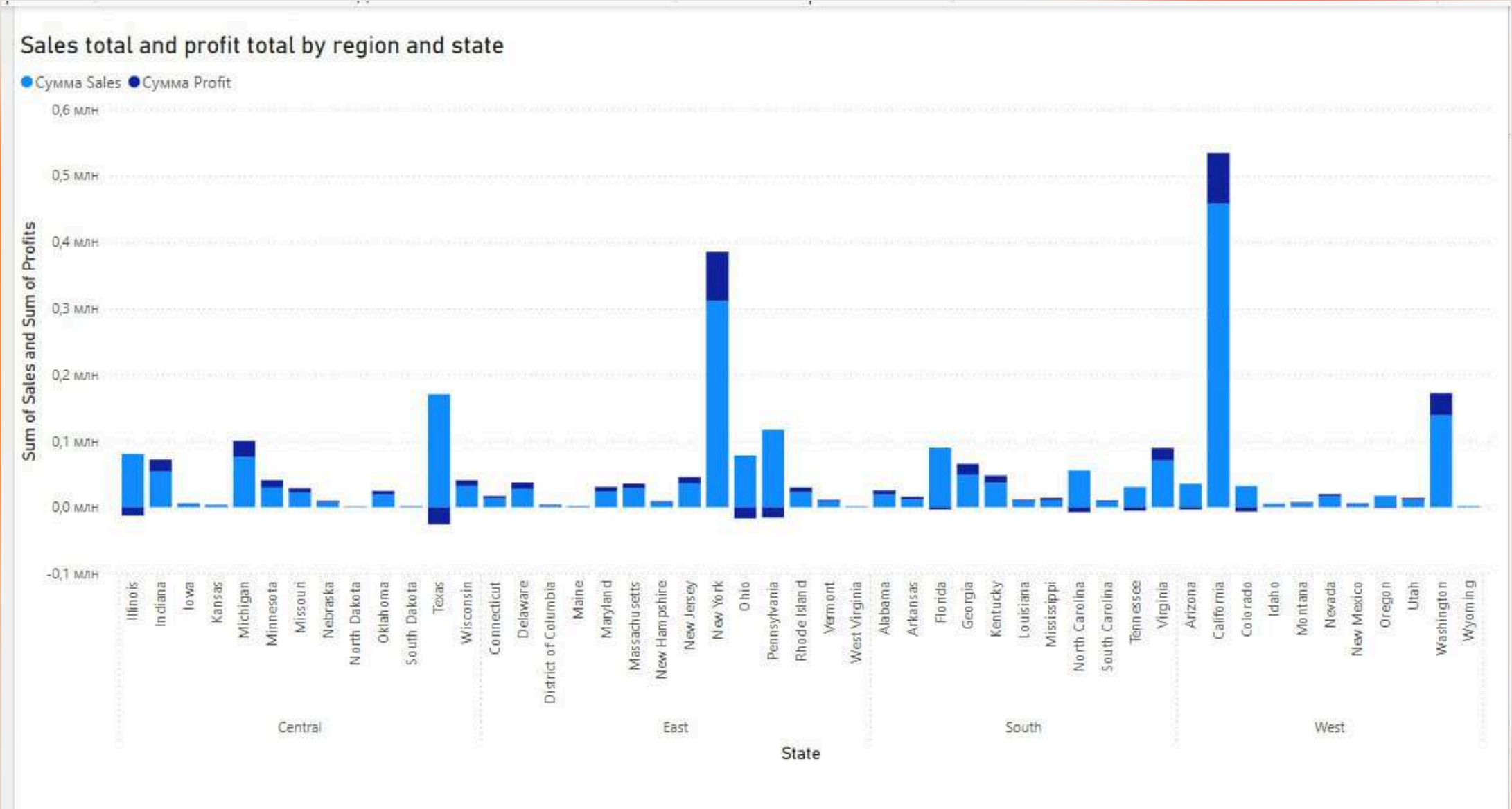
Profit amount by category and sub-category



SUB-CATEGORIES WITHIN THE TECHNOLOGY CATEGORY (E.G., COPIERS, PHONES, ACCESSORIES) ARE SIGNIFICANTLY MORE PROFITABLE THAN SUB-CATEGORIES WITHIN THE FURNITURE CATEGORY (E.G., TABLES, BOOKCASES), WITH TABLES SHOWING A CONSISTENT LOSS.

- RECOMMENDATIONS
- BOOST TECH & HIGH-PERFORMING OFFICE SUPPLIES
 - REVIEW/OPTIMIZE FURNITURE OFFERINGS
 - INVESTIGATE COST STRUCTURE FOR LOW PERFORMERS

Sales total and profit total by region and state

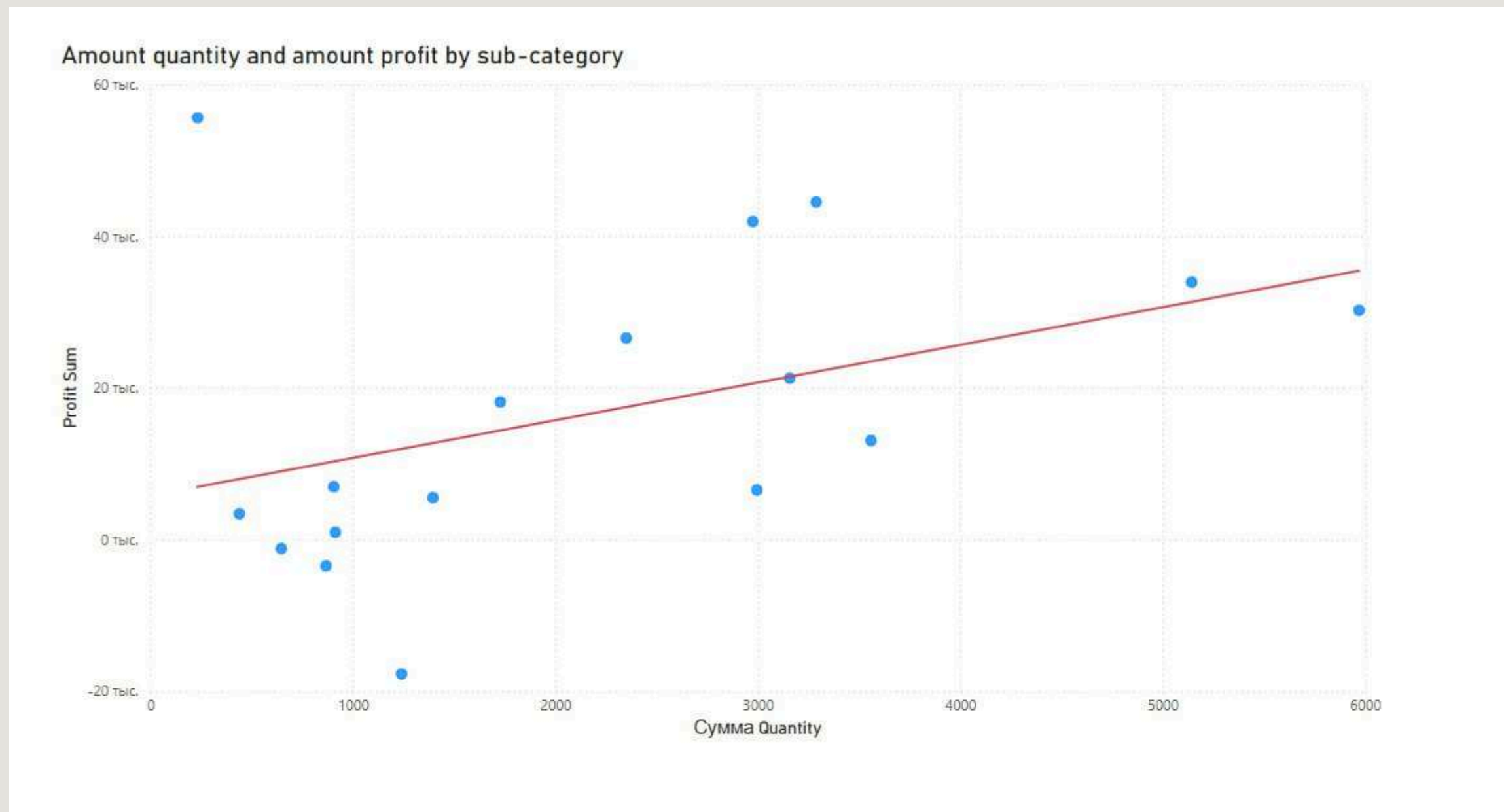


STATES WITH THE HIGHEST SALES VOLUMES, SUCH AS CALIFORNIA AND NEW YORK, ALSO GENERATE THE HIGHEST PROFITS, WHILE SOME STATES LIKE TEXAS AND ILLINOIS, DESPITE HIGH SALES, MAY STILL RESULT IN NET LOSSES OR VERY LOW PROFIT MARGINS.

RECOMMENDATIONS

1. **PRIORITIZE HIGH-PERFORMERS:** FOCUS RESOURCES ON MICHIGAN AND CALIFORNIA.
2. **ADDRESS PROFIT ISSUES:** INVESTIGATE AND RESOLVE ISSUES IN NEW YORK AND TEXAS.
3. **TAILOR REGIONAL STRATEGIES:** OPTIMIZE STRATEGIES BASED ON REGIONAL PERFORMANCE METRICS.

Amount quantity and amount profit by sub-category



THERE IS A POSITIVE CORRELATION BETWEEN THE QUANTITY OF ITEMS SOLD AND THE TOTAL PROFIT ACROSS SUB-CATEGORIES.

THE UPWARD TREND LINE SUGGESTS THAT AS THE QUANTITY SOLD INCREASES, PROFIT GENERALLY INCREASES.

HOWEVER, THERE ARE CLEAR OUTLIERS WHERE HIGH QUANTITY DOESN'T TRANSLATE INTO HIGH PROFIT, AND VICE VERSA.

Top 10 products, bottom 10 products

Top 10 Products

Sub-Category	Cymma Profit
Envelopes	6 964,18
Furnishings	13 052,72
Appliances	18 138,01
Storage	21 278,83
Chairs	26 567,13
Binders	30 228,00
Paper	33 944,24
Accessories	41 936,64
Phones	44 515,73
Copiers	55 617,82
Total	292 243,29

Bottom 10 Products

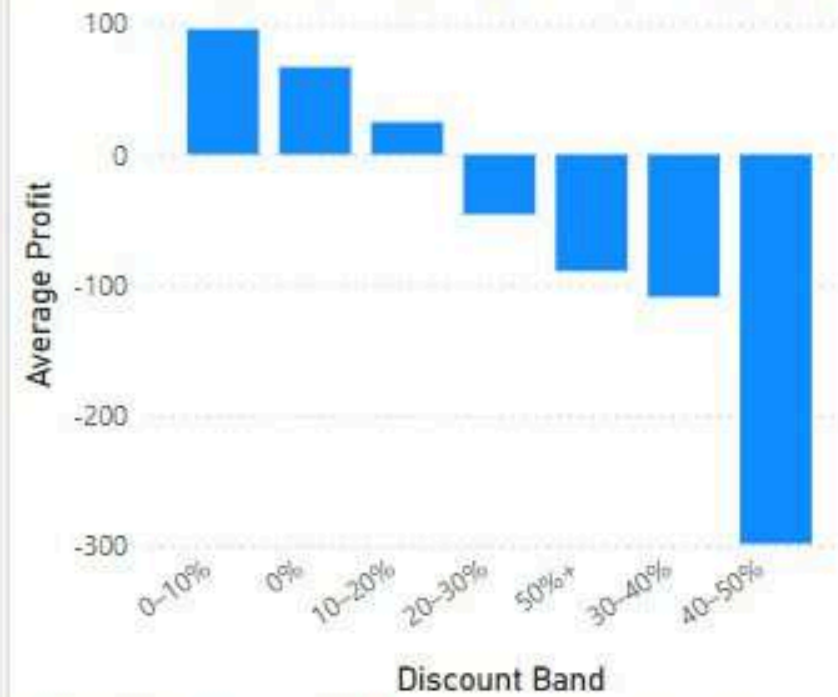
Sub-Category	Cymma Profit
Appliances	18 138,01
Furnishings	13 052,72
Envelopes	6 964,18
Art	6 524,61
Labels	5 526,38
Machines	3 384,76
Fasteners	949,52
Supplies	-1 189,10
Bookcases	-3 472,56
Tables	-17 725,48
Total	32 153,04

KEY INSIGHTS

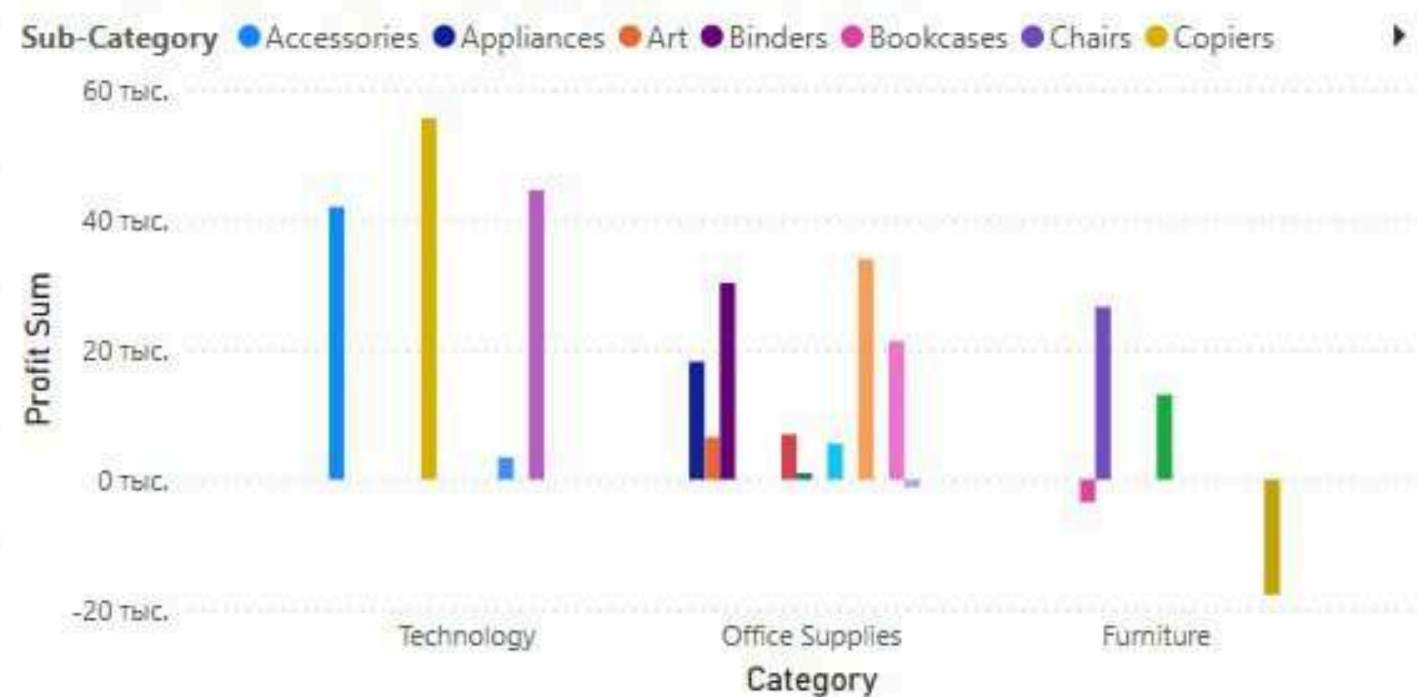
- TOP PRODUCTS:
 - HIGHEST PROFIT: COPIERS (55,258)
 - OTHER HIGH-PERFORMERS: PHONES (45,416), ACCESSORIES (44,911)
- BOTTOM PRODUCTS:
 - LOWEST PROFIT: TABLES (-17,739)
 - OTHER LOW-PERFORMERS: BOOKCASES (-13,775), SUPPLIES (-11,890)

Dashboard

Average profit value by discount band



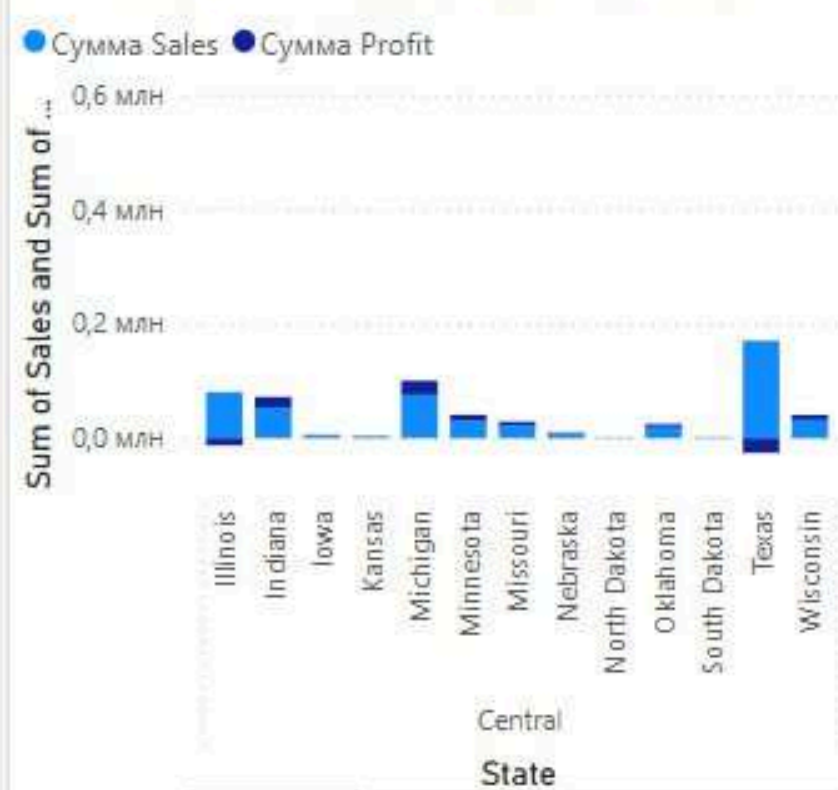
Profit amount by category and sub-category



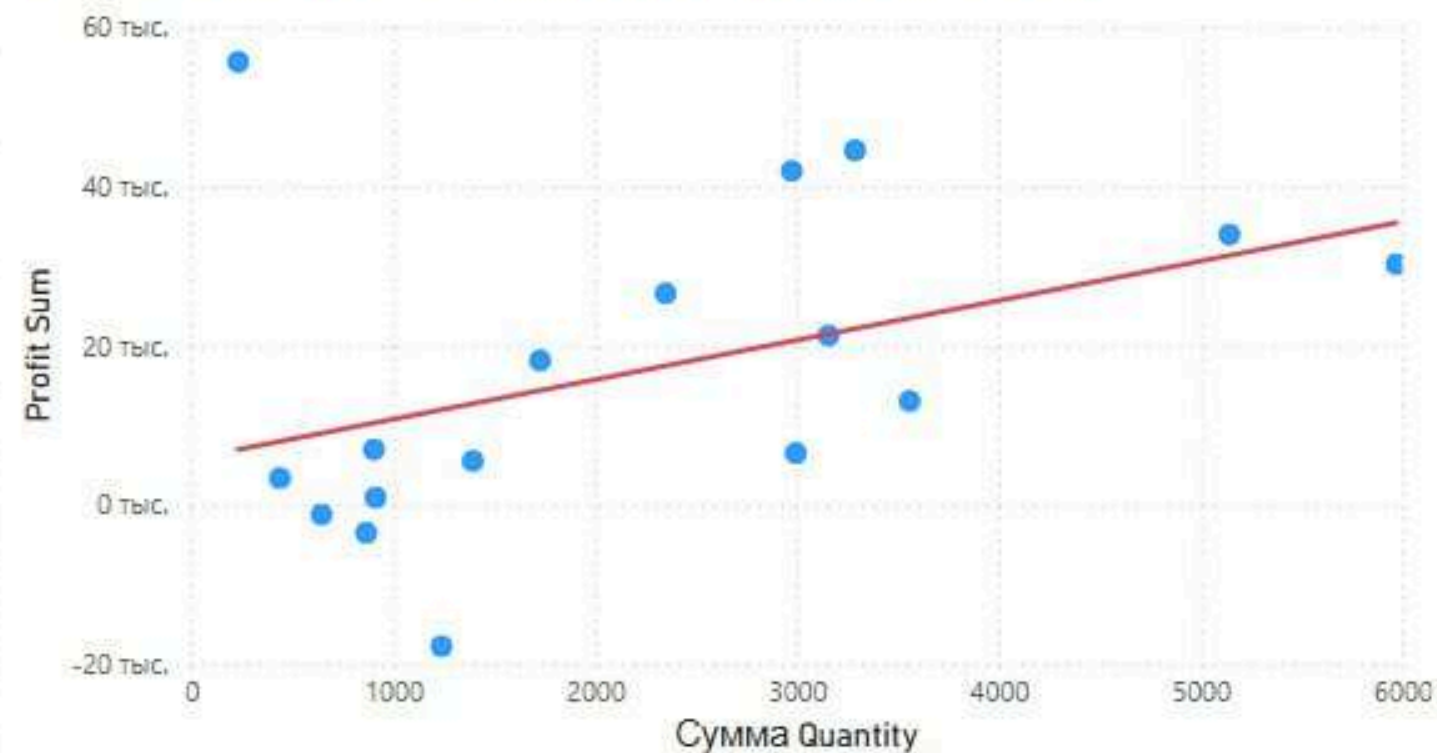
Top 10 Products

Sub-Category	Сумма Profit
Envelopes	6 964,18
Furnishings	13 052,72
Appliances	18 138,01
Storage	21 278,83
Chairs	26 567,13
Binders	30 228,00
Paper	33 944,24
Accessories	41 936,64
Phones	44 515,73
Copiers	55 617,82
Total	292 243,29

Sales total and profit total by region and state



Amount quantity and amount profit by sub-category



Bottom 10 Products

Sub-Category	Сумма Profit
Appliances	18 138,01
Furnishings	13 052,72
Envelopes	6 964,18
Art	6 524,61
Labels	5 526,38
Machines	3 384,76
Fasteners	949,52
Supplies	-1 189,10
Bookcases	-3 472,56
Tables	-17 725,48
Total	32 153,04

Strategic Recommendations

Average profit value by discount band

- IF WE LOOK TO THIS DIAGRAM WE SEE THAT:
- PROFITABILITY SHARPLY DECLINES AS DISCOUNT LEVELS INCREASE.
 - DISCOUNTS EXCEEDING 20% ARE TYPICALLY ASSOCIATED WITH NEGATIVE AVERAGE PROFITS.
 - THE 40-50% DISCOUNT BAND CONSISTENTLY RESULTS IN THE MOST SIGNIFICANT LOSSES.

- RECOMMENDATIONS:
- CAP DISCOUNTS AT A MAXIMUM OF 20% FOR MOST PRODUCT CATEGORIES.
 - IMPLEMENT A TARGETED, DATA-DRIVEN DISCOUNTING STRATEGY BASED ON CUSTOMER SEGMENTATION AND PRODUCT PERFORMANCE.
 - MONITOR PROFIT MARGINS CLOSELY WHEN RUNNING PROMOTIONAL CAMPAIGNS.

- IF WE LOOK TO SECOND DIAGRAM WE SEE THAT,UNDERPERFORMING SUB-CATEGORIES INCLUDE:
- TABLES (SIGNIFICANT NET LOSS)
 - BOOKCASES
 - SUPPLIES

- RECOMMENDATIONS:
- CONDUCT A COST STRUCTURE ANALYSIS TO IDENTIFY CAUSES OF UNPROFITABILITY (E.G., SHIPPING, STORAGE, PROCUREMENT COSTS).
 - REDUCE STOCK LEVELS OR CONSIDER DISCONTINUING PRODUCTS WITH PERSISTENTLY NEGATIVE MARGINS.
 - AVOID DISCOUNTING UNPROFITABLE PRODUCTS, AS DEEPER PRICE CUTS TEND TO AMPLIFY LOSSES.

Profit amount by category and sub-category

Sales total and profit total by region and state

- STRONG-PERFORMING REGIONS INCLUDE:
- TEXAS
 - MICHIGAN
 - INDIANA
- UNDERPERFORMING REGIONS INCLUDE:
- SOUTH DAKOTA
 - NORTH DAKOTA
 - NEBRASKA
- RECOMMENDATIONS:
- ALLOCATE ADDITIONAL MARKETING AND SALES RESOURCES TO HIGH-PERFORMING REGIONS TO MAXIMIZE RETURN ON INVESTMENT.
 - INVESTIGATE THE OPERATIONAL INEFFICIENCIES OR DEMAND ISSUES IN UNDERPERFORMING STATES.
 - CONSIDER REGIONAL PRICING OR DELIVERY STRATEGIES TO IMPROVE PROFITABILITY IN LESS SUCCESSFUL AREAS.

Overall



- IF WE TALK ABOUT PRODUCT,WE SEE THAT TOP-PERFORMING SUB-CATEGORIES INCLUDE:
- COPIERS
 - PHONES
 - ACCESSORIES
 - PAPER
 - BINDERS
- RECOMMENDATIONS:
- INCREASE INVENTORY AND MARKETING INVESTMENTS IN THESE PRODUCT LINES.
 - EXPLORE POTENTIAL FOR STRATEGIC PRICE INCREASES TO MAXIMIZE MARGINS WITHOUT HARMING DEMAND.
 - CONSIDER BUNDLING THESE HIGH-PERFORMING PRODUCTS WITH SLOWER-MOVING ITEMS TO DRIVE OVERALL SALES.

Top 10 products, bottom 10 products and Amount quantity and amount profit by sub-category

TO IMPROVE PROFITABILITY, THE BUSINESS SHOULD FOCUS ON TOP-PERFORMING PRODUCTS LIKE COPIERS AND PHONES, REDUCE OR ELIMINATE UNPROFITABLE ITEMS SUCH AS TABLES, AND LIMIT DISCOUNTS TO NO MORE THAN 20%. RESOURCES SHOULD BE CONCENTRATED IN HIGH-PERFORMING REGIONS, AND INVENTORY SHOULD BE MANAGED BASED ON PROFIT, NOT JUST VOLUME. THESE STEPS WILL LEAD TO STRONGER MARGINS AND MORE EFFICIENT OPERATIONS.

thank you