

FINAL PROJECT

DIVE TO WISH SUMMER

TECH-DA104-4 Python Programming
for Data Analytics
GROUP 4: Willian, Yui, Gui
2025/06/12





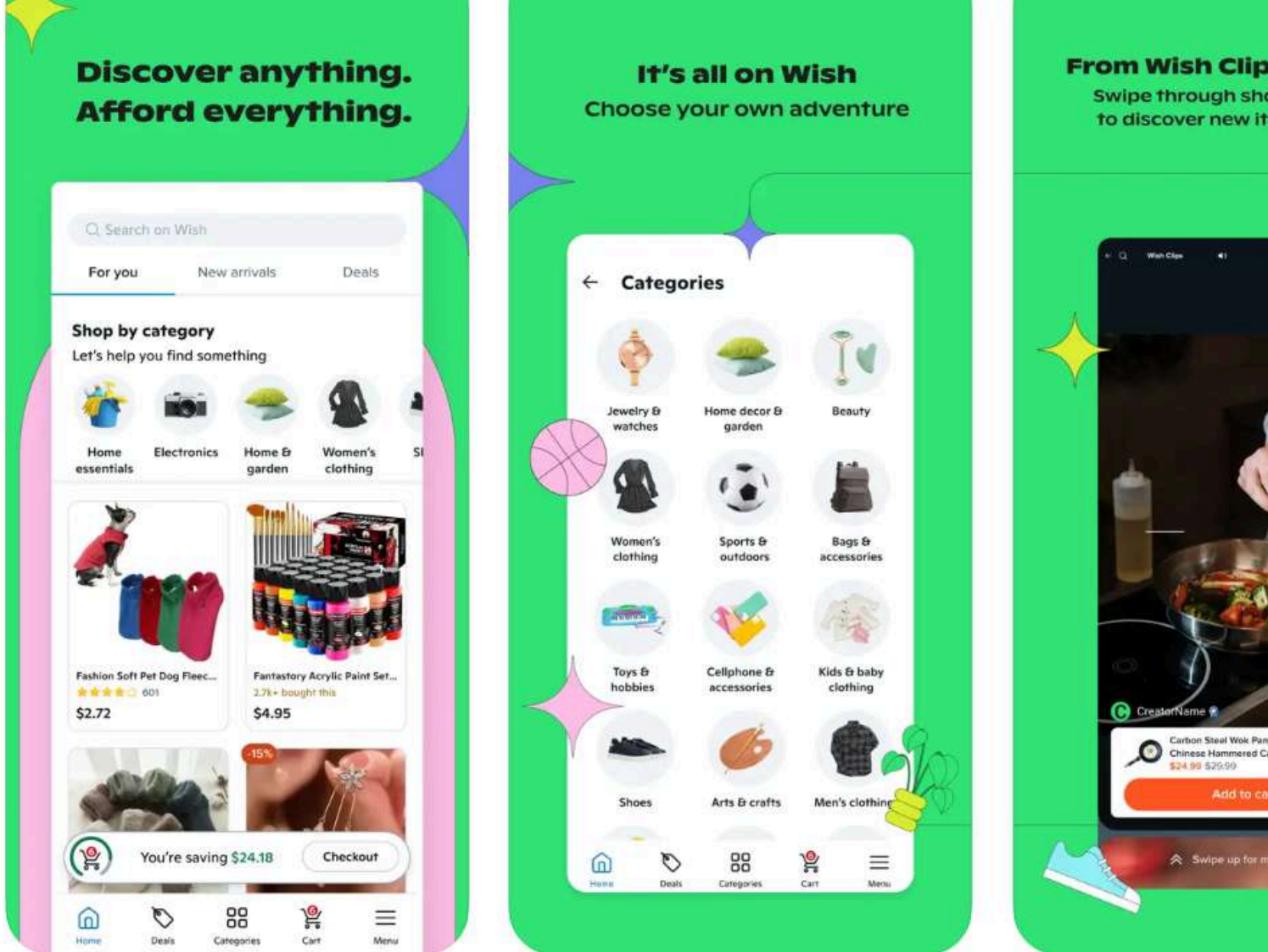
AGENDA

1. ABOUT **WISH**
2. DATA CLEANING
3. OVERVIEW **WISH**
3. SALES DATA ANALYSIS
4. RECOMMENDATIONS

1. ABOUT wish

Wish is an online shopping platform that offers affordable items for your home and life needs.

Launch date	Jul 4, 2010
HQ	San Francisco, California
People	Joe Yan (CEO)
Business type	Subsidiary
Industry	Ecommerce
Owner	Qoo10



Wish Built An \$11 Billion Business On Insanely Cheap Shipping

(Forbs, 2020/07/30)





DATA CLEANING

With Python
Data Base :
Subject_1-Summer
products.csv
[First] → [After]
Column: 43 → 37
Rows: 1573 → 1341



HOW TO STEP

**DATA
CLEANING**

1. Got to know the data
2. Deleted the no-need columns
3. Deleted duplicates productID
4. Summarized gender
5. Filled in the missing colours
6. Categorize clothes

wish

OVERVIEW

WISH

Analyze the wish based on the data after data cleaning.

Merchant

Products

Color

Gender



OVERVIEW Wish

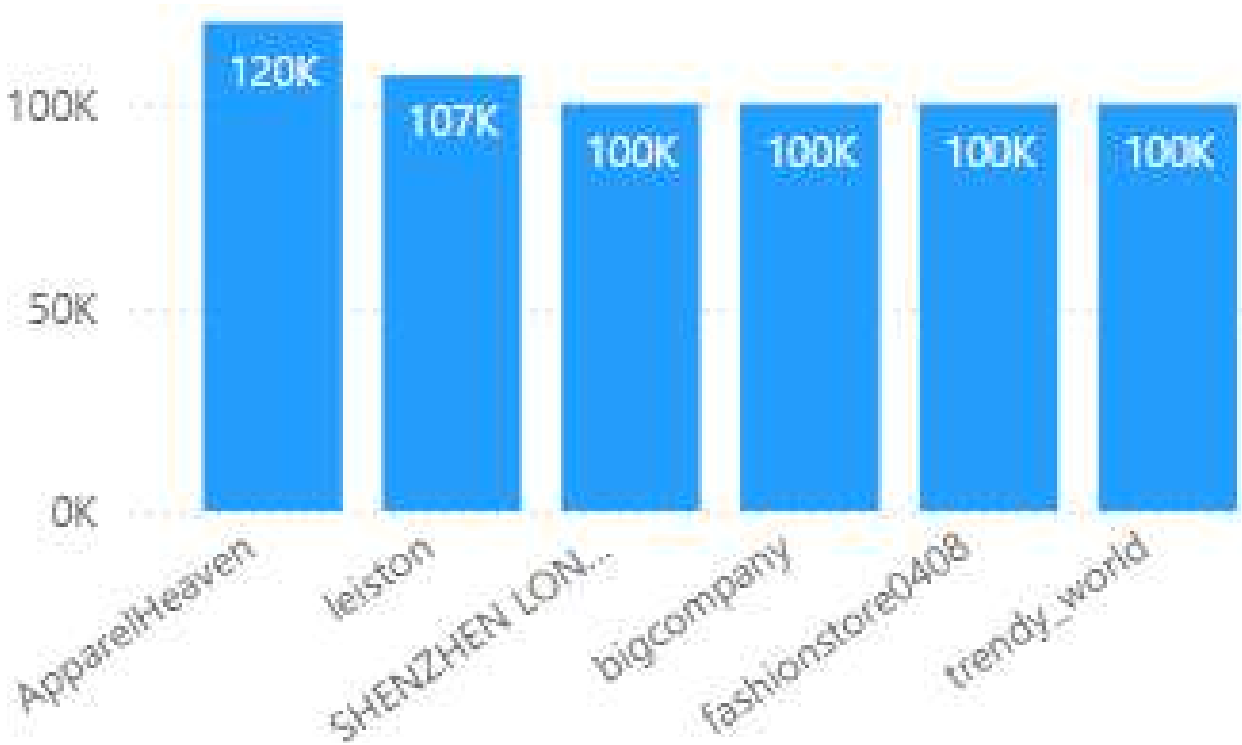
Revenue
52,3M
Units Sold
6,4M
Merchant Count
959
Countries Shipped To
140

Data Date: 2020/08/01

Heat Map of Origin Countries



Top 5 Merchant

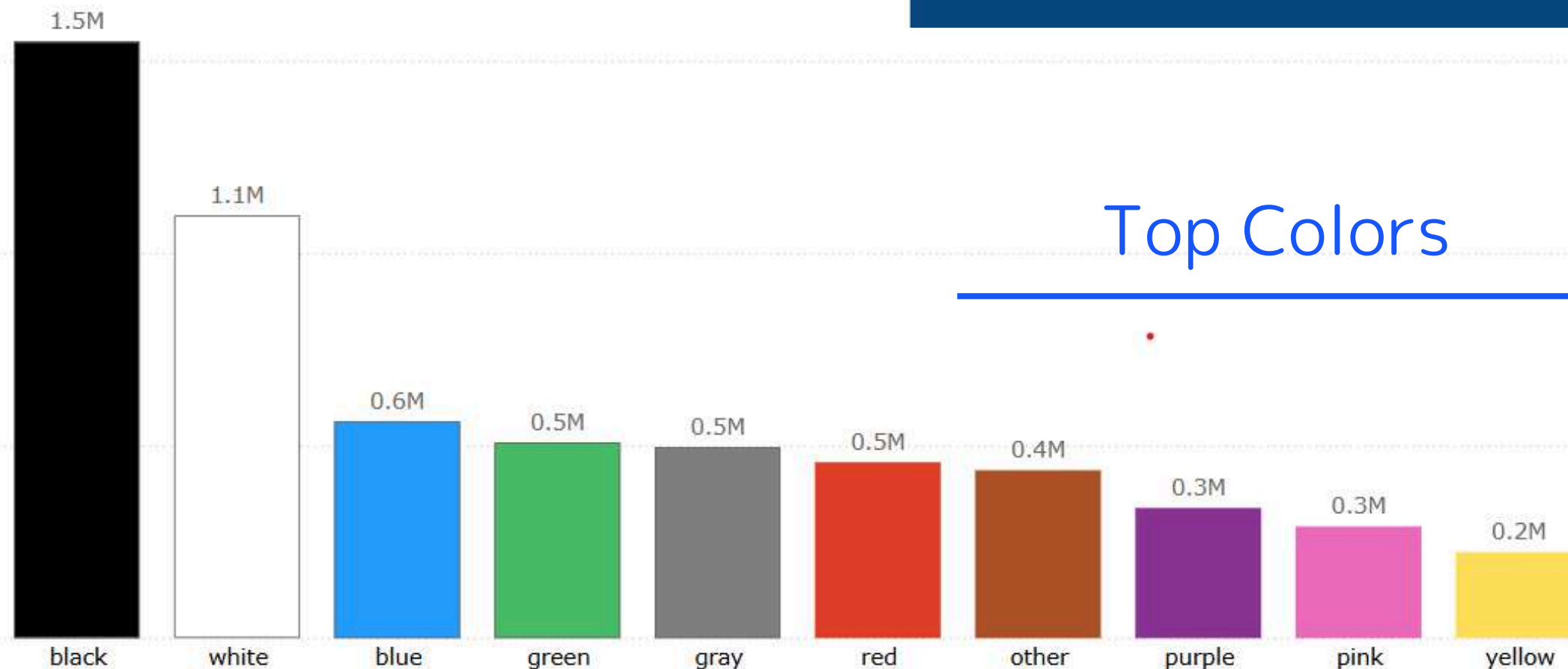


Average Merchant Ratings



OVERVIEW Wish

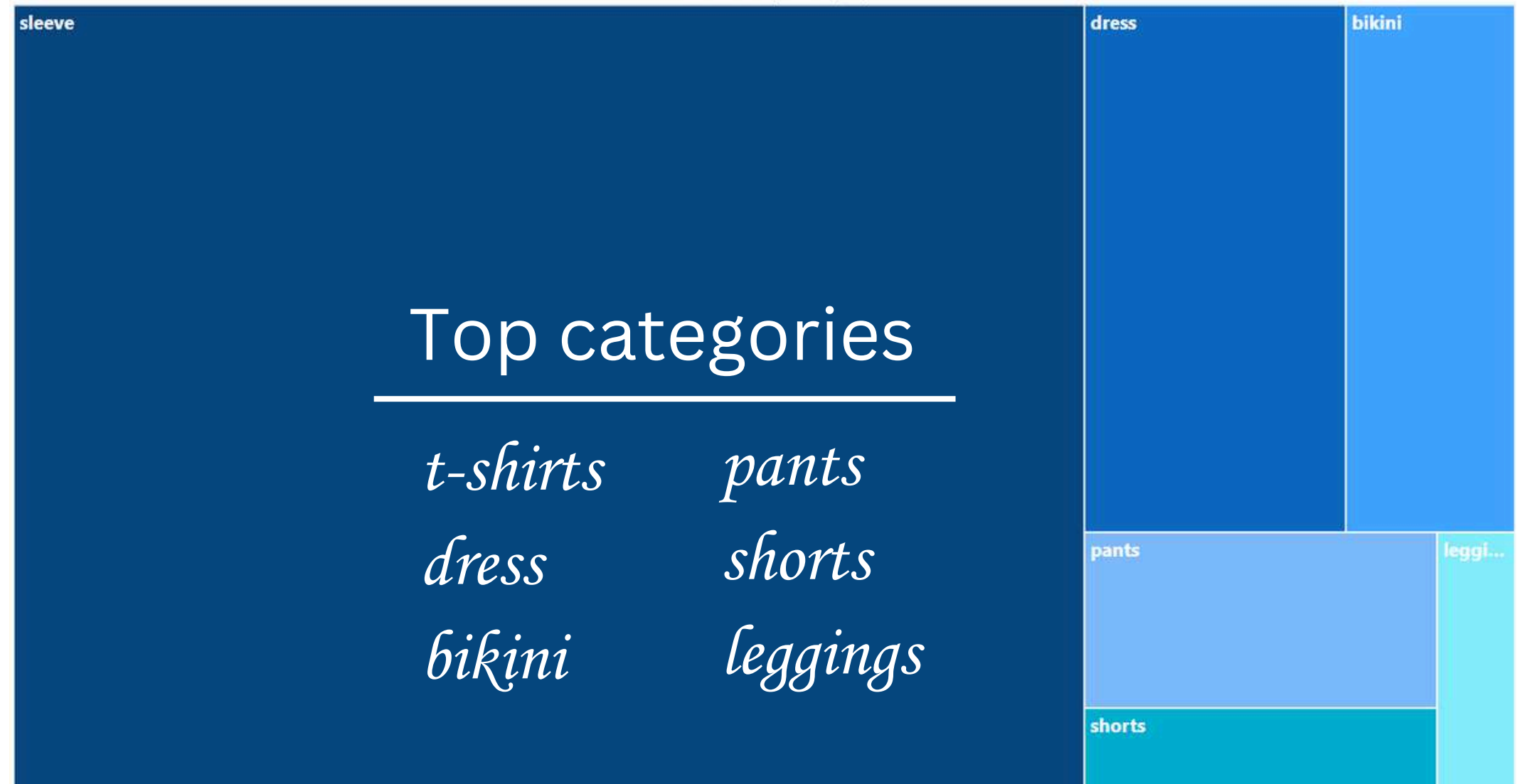
Average Rating Products



Top Colors

Top categories

t-shirts *pants*
dress *shorts*
bikini *leggings*



Color aggregation:
from 101 to 14 colors
e.g.
coffee -> brown
floral -> multicolor
grey -> gray

OVERVIEW wish

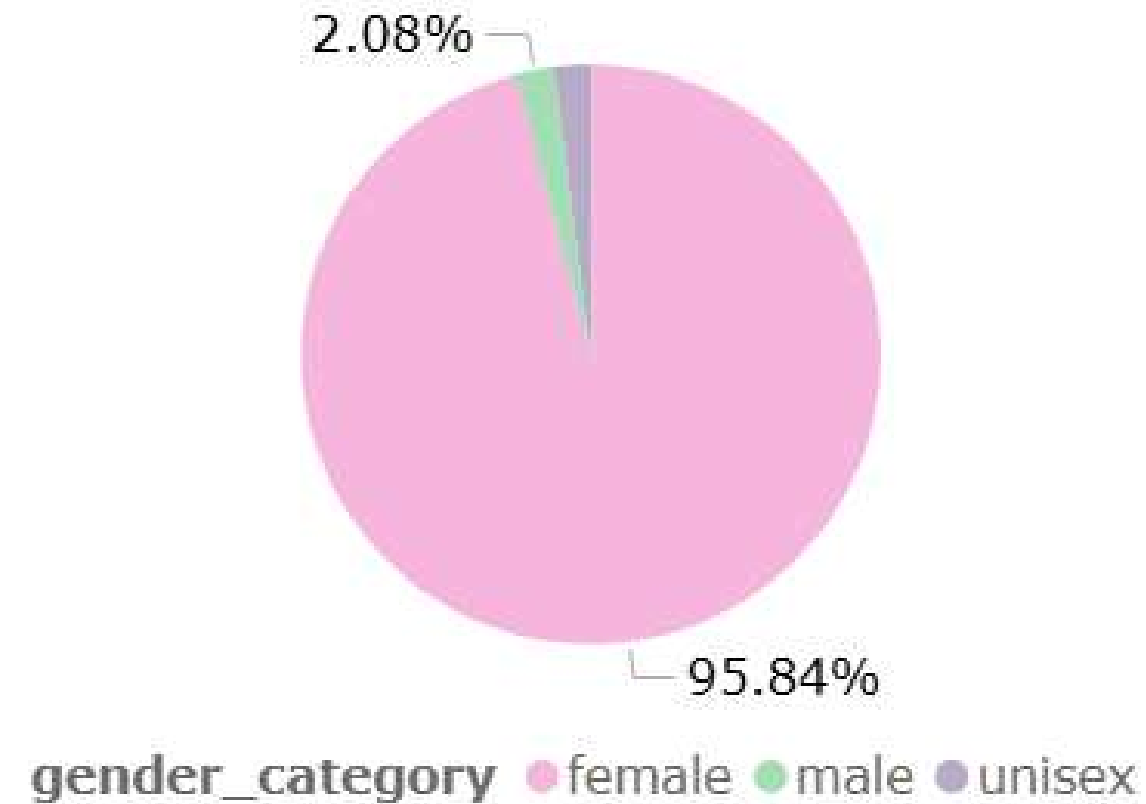
Gender of customer

Female: 6,190,497 (96%)

Male: 134,560 (2%)

Unisex: 134,452 (2%)

Units Sold By Gender



Gender aggregation:
e.g.
women, girls,
'lady, ladies etc
----->>> female
gentlemen, male,
masculine etc
-> male

Revenue by Category and Gender



SALES DATA ANALYSIS

1. About “ad_boosts” and “urgency_banner”
2. What kind of category sell the most?
3. Is the shipping cost related to sales?

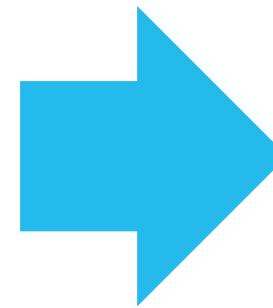


1. ABOUT AD_BOOSTS AND URGENCY_BANNER

SALES DATA ANALYSIS

They have 2 sales methods.

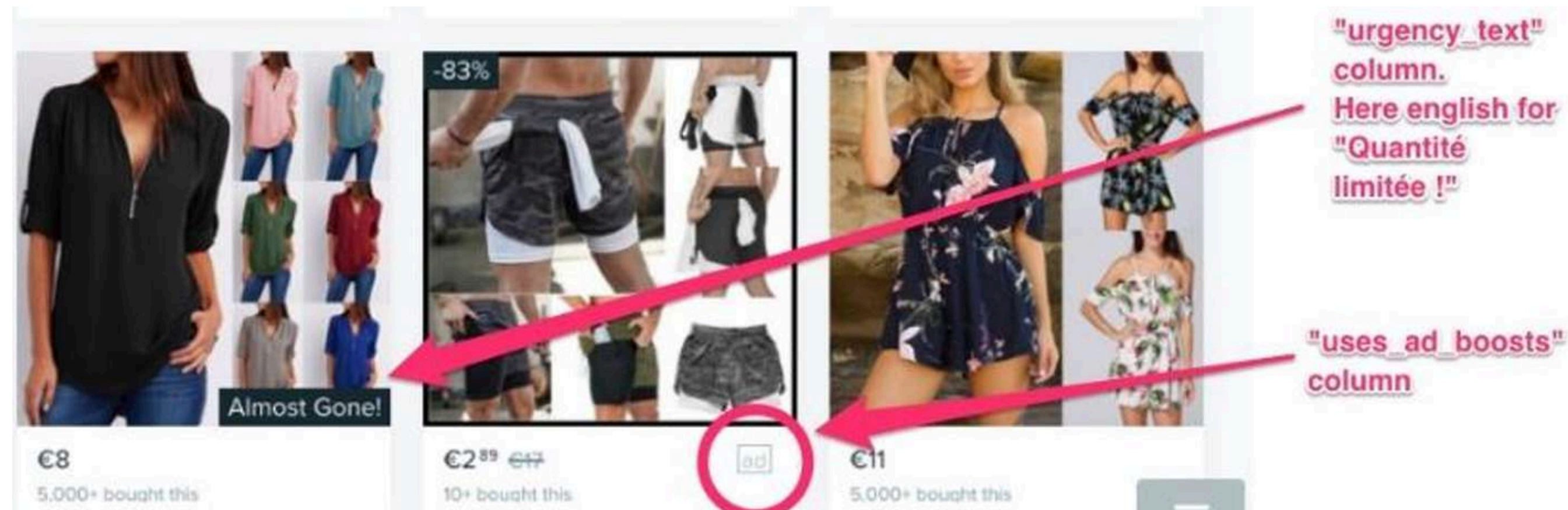
1. `uses_ad_boosts`
2. `urgency_banner`



FOR merchant

What are these effects?

Compare the amount of sales



wish

1. ABOUT AD_BOOSTS AND URGENCY_BANNER

SALES DATA ANALYSIS

HALF THE PRODUCTS USE THIS METHOD.

ad_boosts	urgency_banner	count
×	×	608
×	○	149
○	×	463
○	○	121

total: 1341 rows

[Pair]

urgency_banner = urgency_text

urgency_text (all 270)

- Quantité limitée ! (269)
- Réduction sur les achats en gros (1)

“Quantité limitée !” = Limited quantities!

“Réduction sur les achats en gros”
= Wholesale discount

How do these affect sales?

1. ABOUT AD_BOOSTS AND URGENCY_BANNER

SALES DATA ANALYSIS

urgency_banner



They can boosting sales!

→ Effective for selling



ad_boosts

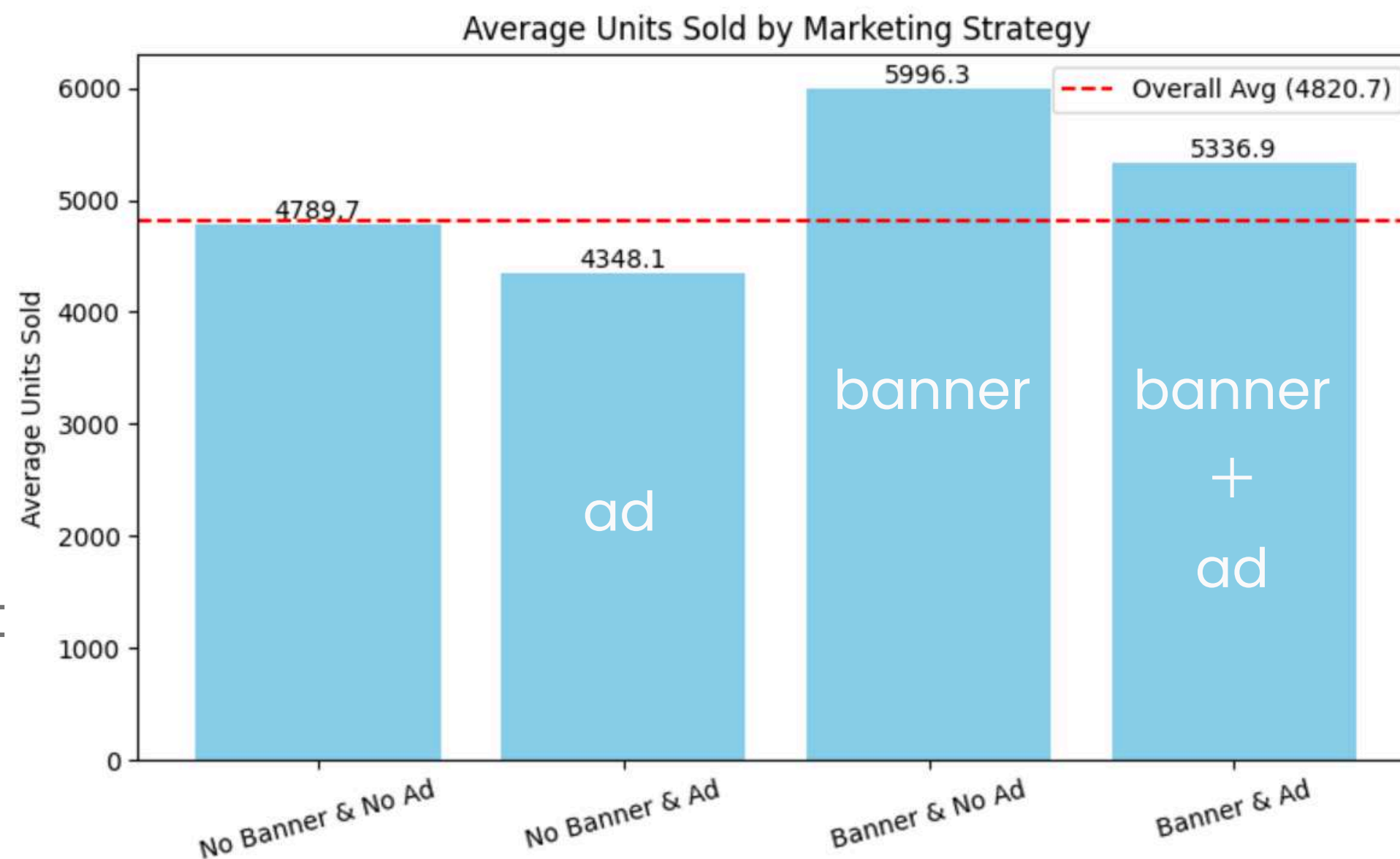


Effects looks bad.

→ Why?

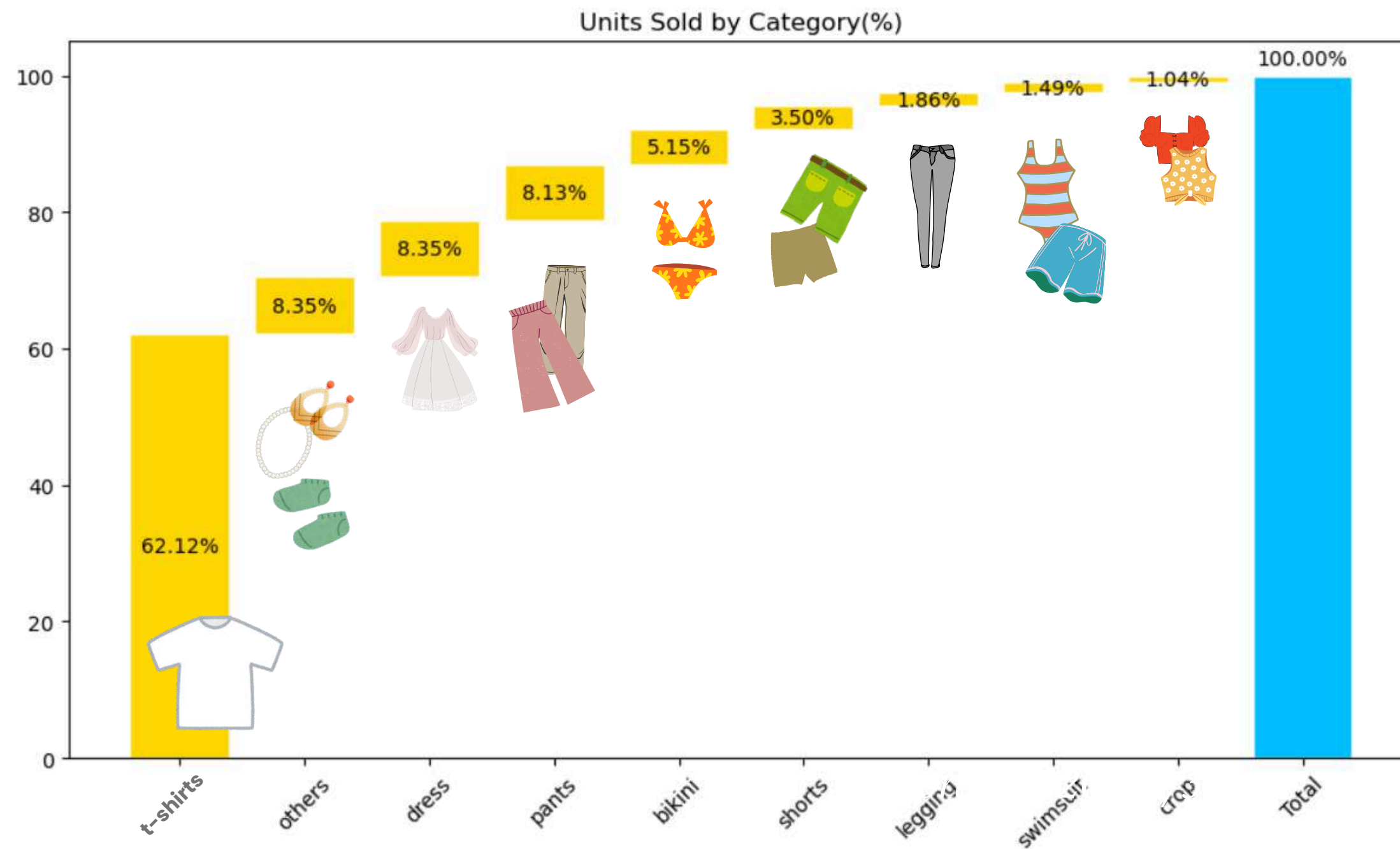
[Prediction] Used for products that don't sell?

(To be displayed on the top page)



2. WHAT KIND OF CATEGORY SELL THE MOST?

SALES DATA ANALYSIS

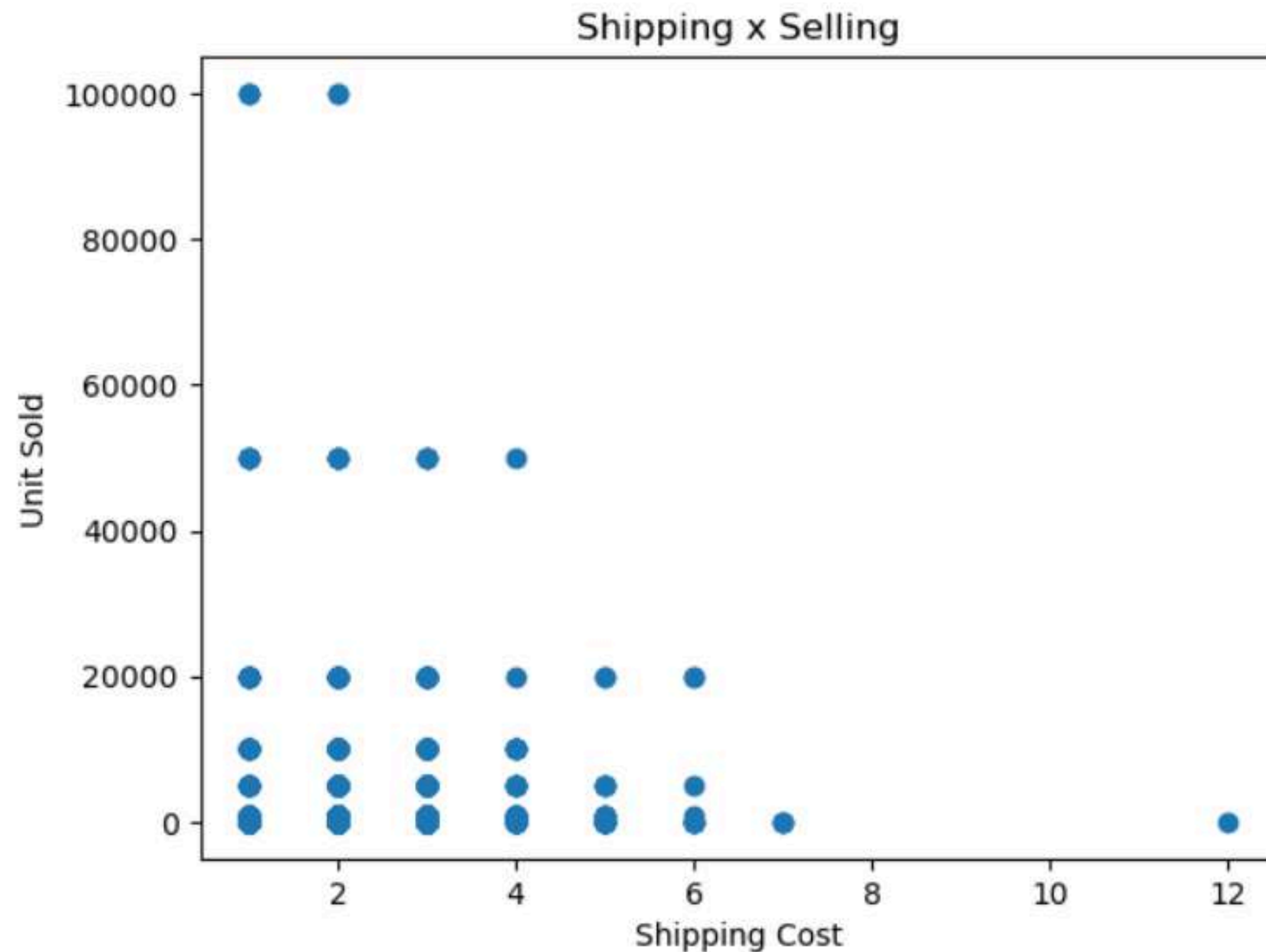


Category t-shirts is the top seller.



3. IS THE SHIPPING COST RELATED TO SALES?

SALES DATA ANALYSIS



Lower shipping costs are strongly associated with higher sales volumes.





RECOMMENDATIONS

Methods for Selling:

- Use urgency_banner & text
- t-shirts category
- Nearby areas with low shipping costs

→ Boosting poor-selling products by adding ads to them

THANK YOU!



TECH-DA104-4
Python Programming for
Data Analytics
GROUP 4: Willian, Yui, Gui

