

# Identifying Factors that drive *Wish* revenue



# Introduction

- ✓ Wish is interested in looking into factors that drives revenue on summer products.
- ✓ This e-commerce employs browsing technologies and eliminate distributor fees for the sellers as they provide a platform to sell directly to customers.
- ✓ As the company is known by its extremely cheap prices, so variables other than price should be part of the analysis.
- ✓ Revenue comes from three activities: Sales fee (15%), logistic services offered to merchants and by selling advertisement on its platform.



# Hypothesis

Some of the biggest driving factors should be:

For sales:

Quality  
Units sold  
Ratings

For Advertisement:

Ratings  
Fees

For Logistic Service:

Location  
Fees



# Data Description

1573 rows, 43 features

Raw data from Kaggle: “Sales of summer clothes in E-commerce Wish” retrieved from <https://www.kaggle.com/jmmvutu/summer-products-and-sales-in-ecommerce-wish>.

Promising variables are:

Binary:

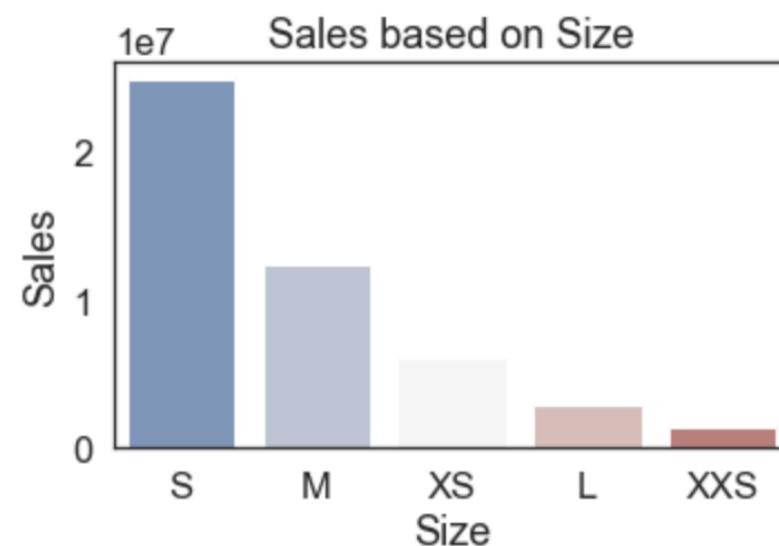
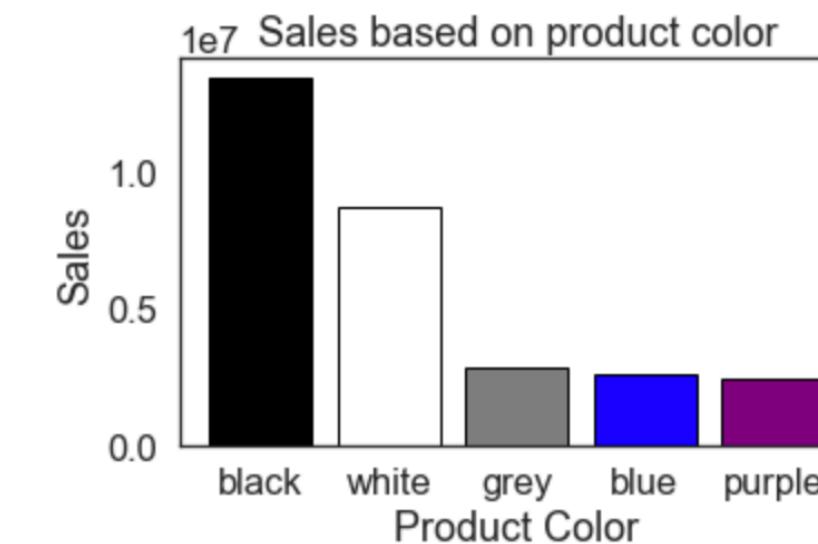
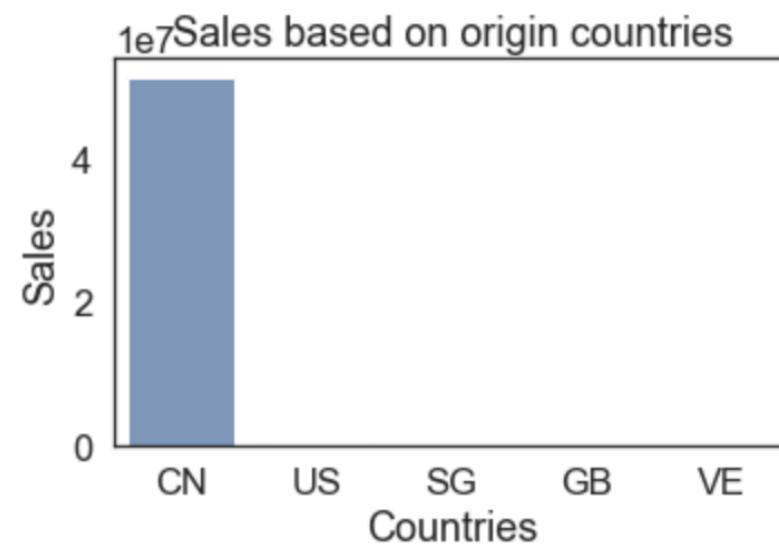
- ✓ Uses ad boosts
- ✓ Badge product quality
- ✓ Merchant has profile picture

Numerical:

- ✓ Units sold
- ✓ Rating count
- ✓ Sales
- ✓ Product rating



# Exploratory Data Analysis



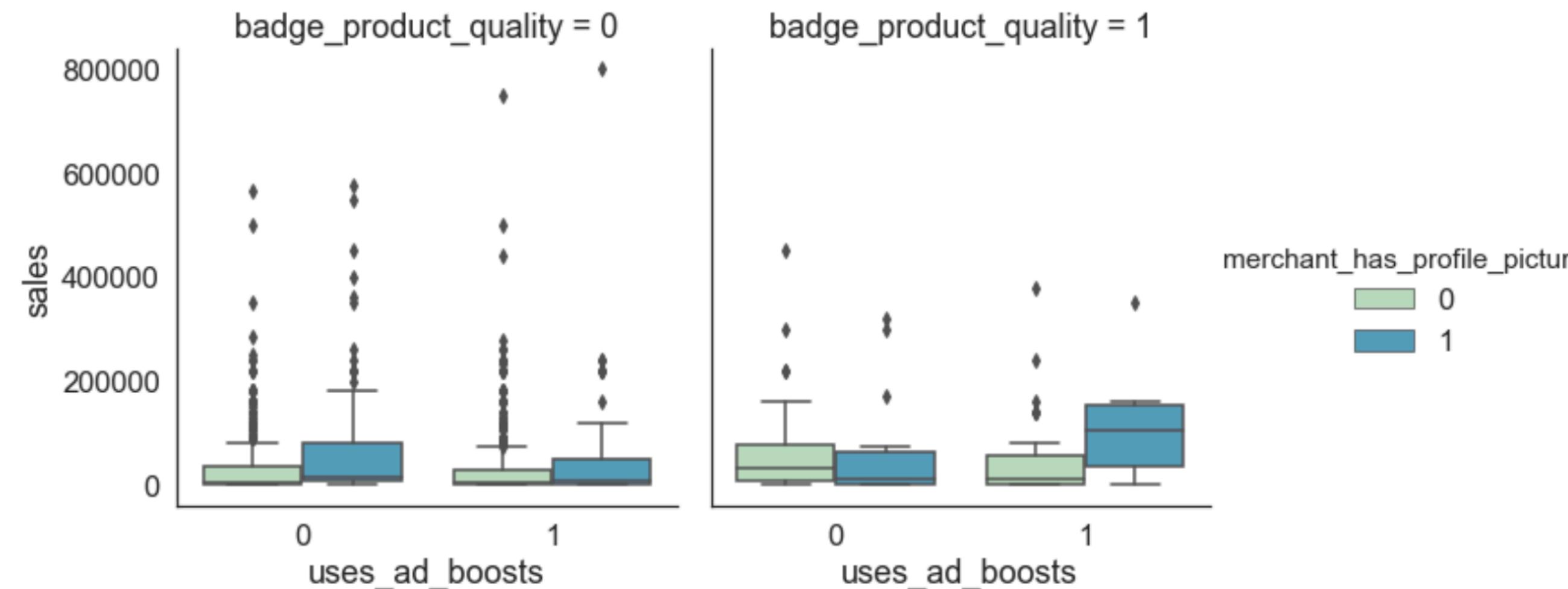
In terms of sales, merchants are from China, and buyers usually buy black products in small size for summer season

Sales have a strong correlation with units sold, and rating count

price	1	0.3	0.05	0.88	0.04	0.15	-0.02	0	0.07	-0.02	-0.01
retail_price	0.3	1	0.03	0.26	0.06	0.07	0.01	0.05	0.03	0.04	0.05
rating	0.05	0.03	1	0.06	0.22	0.06	0.06	0.07	0.19	0.03	0.06
shipping_option_price	0.88	0.26	0.06	1	0.04	0.12	-0.03	-0.01	0.08	-0.05	-0.02
merchant_rating	0.04	0.06	0.22	0.04	1	0.15	0.13	0.15	0.21	0.12	0.16
sales	0.15	0.07	0.06	0.12	0.15	1	0.91	0.83	0.07	-0.03	0.2
units_sold	-0.02	0.01	0.06	-0.03	0.13	0.91	1	0.9	0.05	-0.02	0.23
rating_count	0	0.05	0.07	-0.01	0.15	0.83	0.9	1	0.07	-0.01	0.22
badges_count	0.07	0.03	0.19	0.08	0.21	0.07	0.05	0.07	1	0.03	0.08
countries_shipped_to	-0.02	0.04	0.03	-0.05	0.12	-0.03	-0.02	-0.01	0.03	1	0.13
merchant_rating_count	-0.01	0.05	0.06	-0.02	0.16	0.2	0.23	0.22	0.08	0.13	1

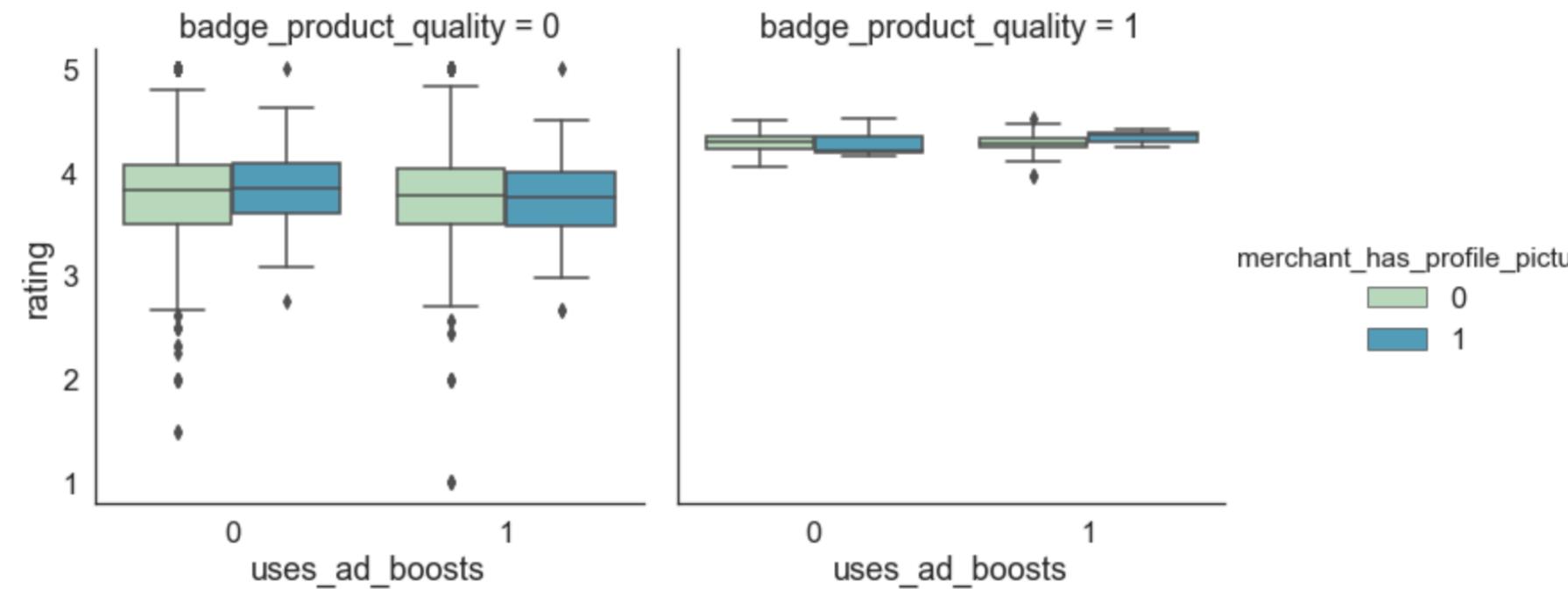
price	
retail_price	
rating	
shipping_option_price	
merchant_rating	
	sales
	units_sold
	rating_count
	badges_count
	countries_shipped_to
	merchant_rating_count

# Exploratory Data Analysis



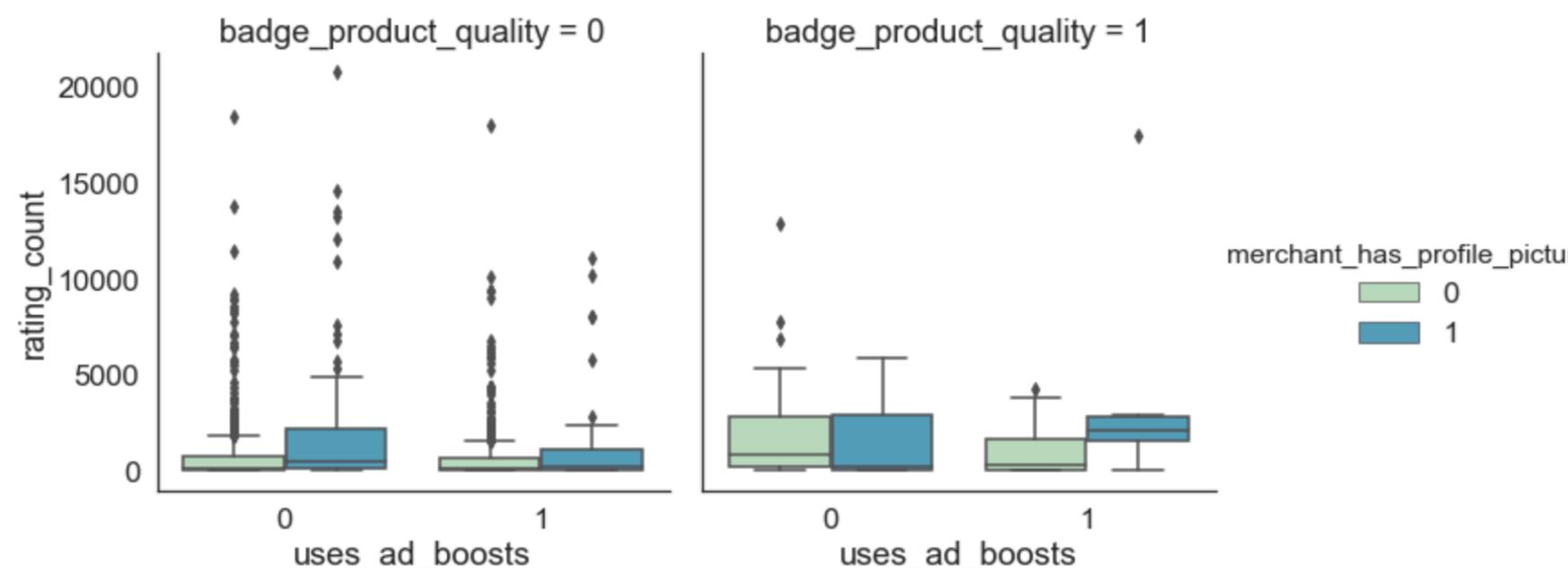
Sales mean is higher when the product has a quality badge, is promoted by using an ad boost and the merchant has a profile picture.

# Exploratory Data Analysis

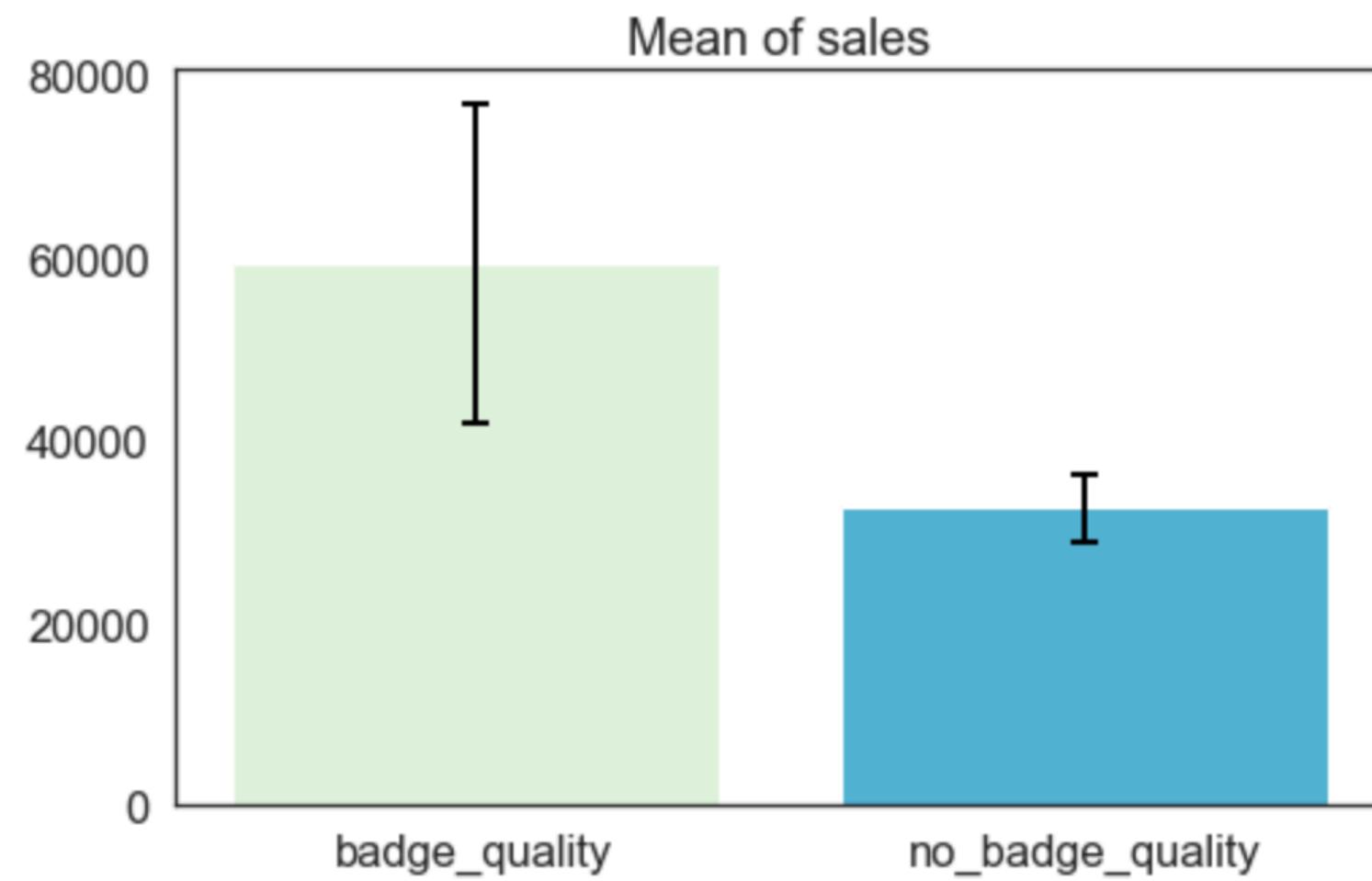


When the product has a **quality badge**, is promoted by an **ad boost** and the merchant has a **profile picture**, product rating and rating Count is higher.

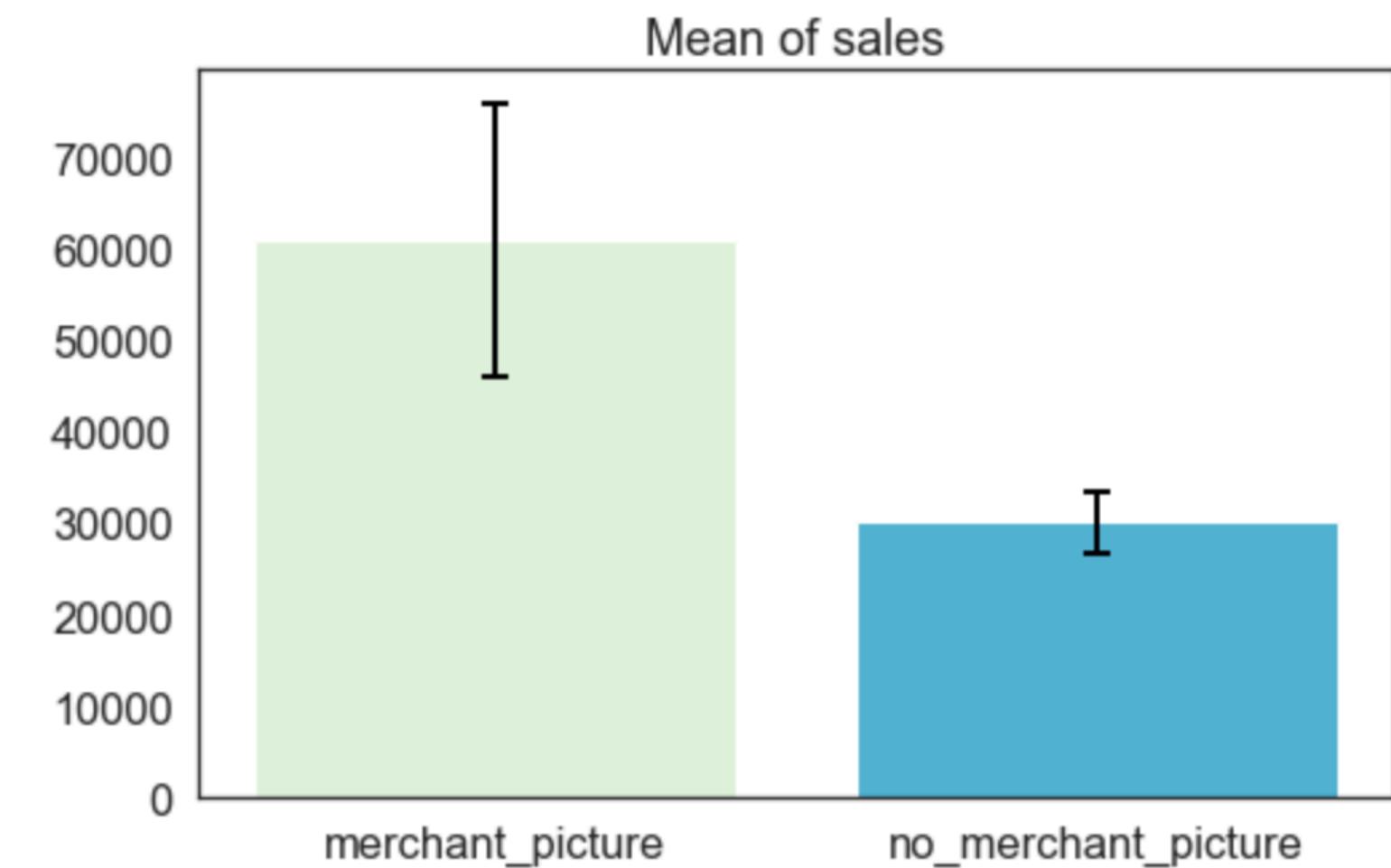
However, rating is not correlated to rating count.



# T-Test on Sales

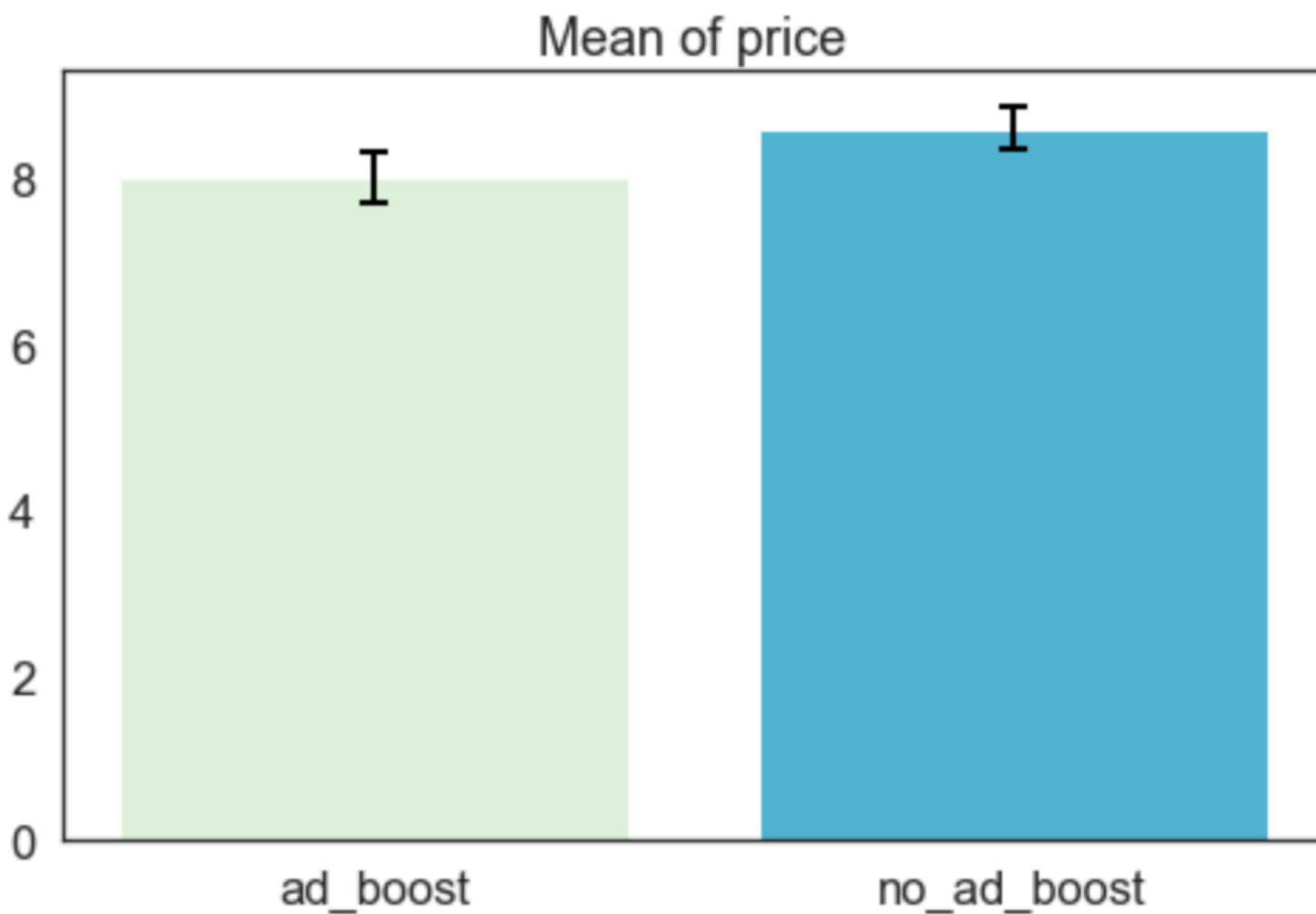


With 95% confidence, sales are between \$9K and \$44K higher for products with a quality badge.

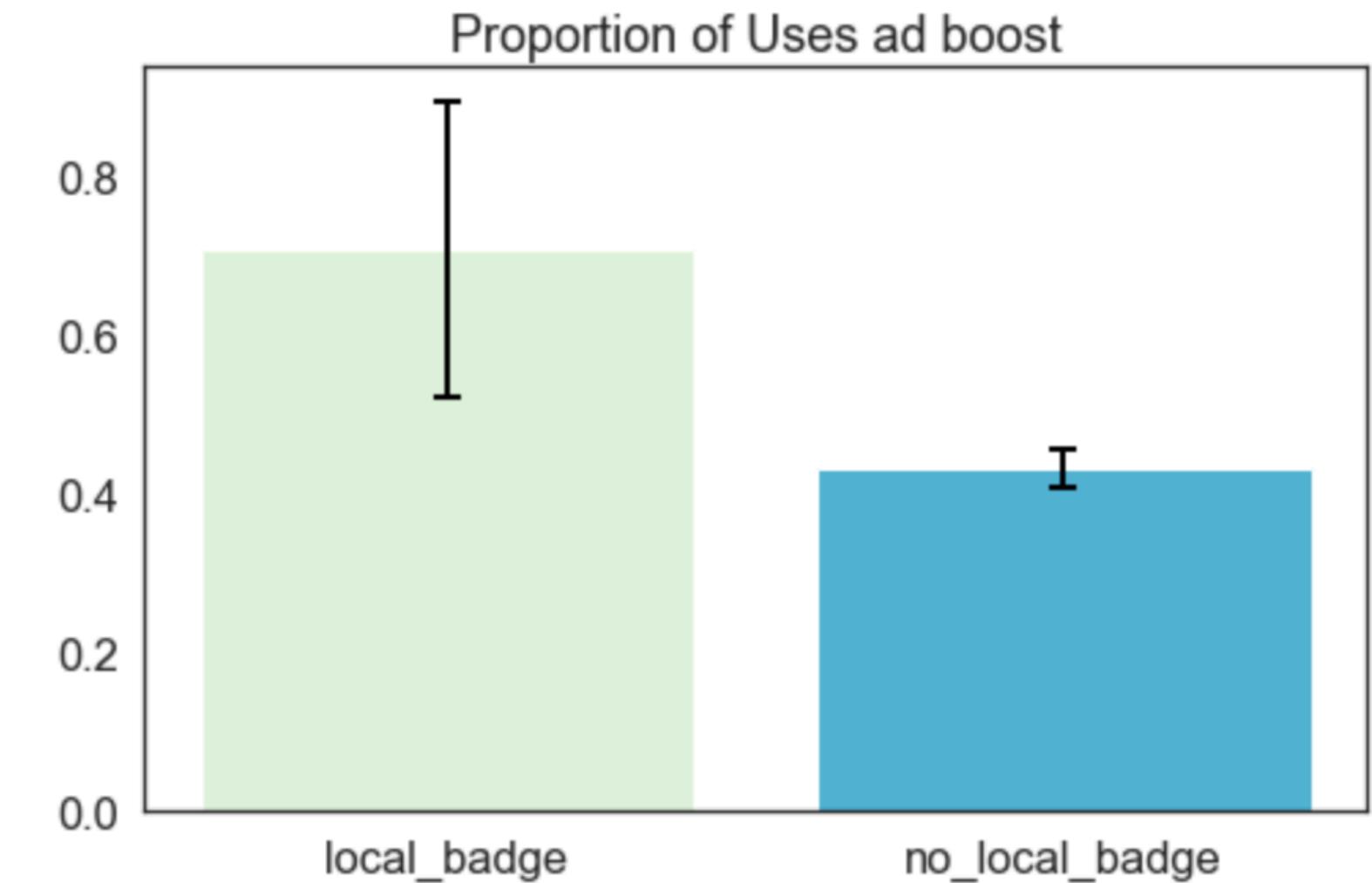


With 95% confidence, sales are between \$15K and \$45K higher for products whose merchant has a profile picture.

# Ad Boost



With 95% confidence, product prices are between \$0.19 and \$0.99 higher for products that are not promoted by ad boost.



\*\*With 95% confidence, ad boost over products is higher between 7% and 48% of local products. Sampling error?

# Takeaways

- ✓ Units sold and rating counts show the strongest relationship with sales.
- ✓ Sales mean is higher when the product has a quality badge, is promoted by using an ad boost and the merchant has a profile picture.
- ✓ Product rating is not product rating count. The first shows a Pearson R of 0.06 when correlated with Sales, the latter has a strong correlation of 0.83.
- ✓ To test hypothesis for this binary variable, it was performed a z-test. However, there wasn't enough data to validate the only experiment that got a p-value lower than 0.05, but it looks like merchant uses ad boosts when products are local.





**Thank you!**