E-commerce:
Product Range Analysis

•••

Lior Kedem Kadish
July 2022

Project objective:
To identify best sellers vs. failed products

Conclusions - Best sellers

The products

4 Best sellers & most gainful products:

- "White hanging heart t-light holder"
- "Jumbo bag red retrospot"
- "Regency cake stand 3 tier"
- "Party bunting"

The sales

- The 10 best sellers accounted for 3% of purchases, and for 7% of the revenue
- They exceeded other products in quantity & price (sold 17 units vs. 10, on average)

The time trend

- Their purchases - but not their revenue - varied by the season, similarly to the general time trend
- The high season was the fall. The peak was November 2019

Conclusions - Clusters

The clusters

- 34% of the products can be clustered as "useless" (with low purchases, low revenue, and long time since purchase)
- 30% more require different levels of attention

The sales

- The "best products" cluster sold more units than the “useless” cluster (10 units vs. 6.8), but its products were cheaper (3.3$ vs. 3.9$)

The time trend

- Besides the "best products”, the rest of the clusters showed few fluctuations over the year

Conclusions - Specific

Jumbo products

- Only 25 “Jumbo” products, yet they accounted for 4% of purchases, and for 6% of the revenue
- They were significantly more likely to be successful than other products - both in terms of purchases and in terms of revenue

Recommendations

General

- Make special offers during winter and summer, to boost sales in times of decline

Clusters

- Consider replacing the products from the "useless" cluster
- Market the "great but not recent enough" and the "potential" clusters
- Make special offers for the "gainful but forgotten" cluster
- Make special offers for high quantity of the "popular but gain little" cluster

Specific

- Increase the selection of the successful jumbo products, for example by adding new designs

When do the sales go up?

The highest amount of purchases - in the fall

Findings

- The peak: November 2019
- The lowest amount of purchases - in the winter. Sales were better in the spring, and lower in the summer

* November 2018 & December 2019 are excluded due to lack of data; April is low in part because of the gap in the data during Easter

The highest revenue - also in the fall

Findings

- Relatively similar to the time trend of the number of purchases
- August a bit better here, while April, May & June a bit worse

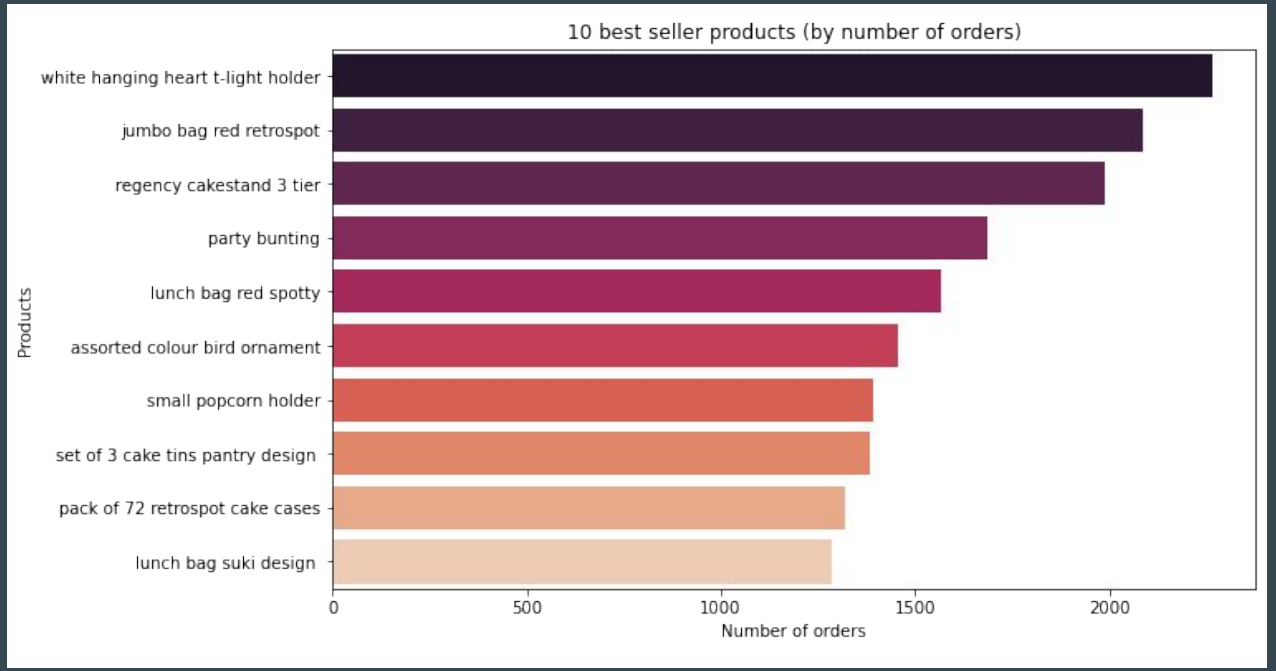
What are the most successful products?

(based on RFM analysis)

The best seller: "white hanging heart t-light holder"

Findings

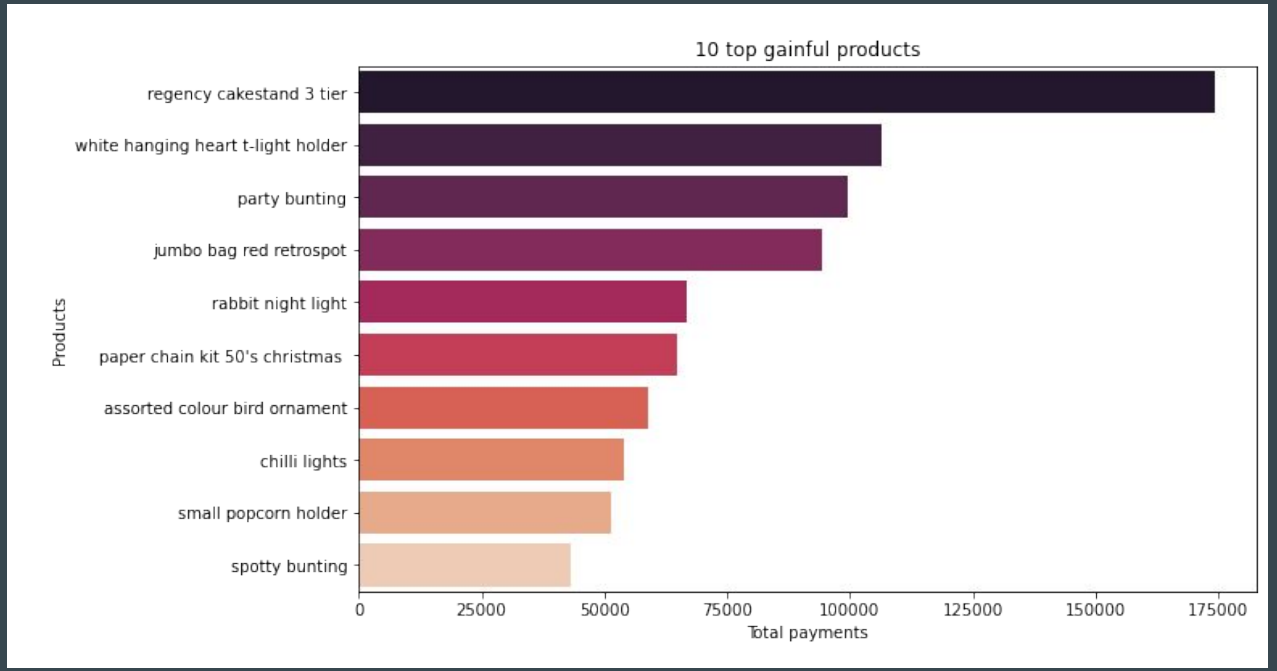
- The best seller was purchased 2,265 times
- The two following products were purchased about 2,000 times
- All were purchased more than 1,200 times



The 4 most gainful products are the 4 best sellers

Findings

- The most gainful product: “regency cake stand 3 tier”
- It gained 174,157$ - substantially higher than the others (4 times more than the 10th)



Failed products

Least popular

127 products
were bought only once

Least recent

82 products
were purchased more than a year before our recent time limit

Least gainful

80 products
gained less than 5$ in total

“Pads to match all cushions” gained the least (close to 0$)

What are the attributes of the best sellers?

The 10 best sellers excelled in revenue and recency

Purchases

They accounted for 3% of the purchases

Revenue

They accounted for 7% of the revenue

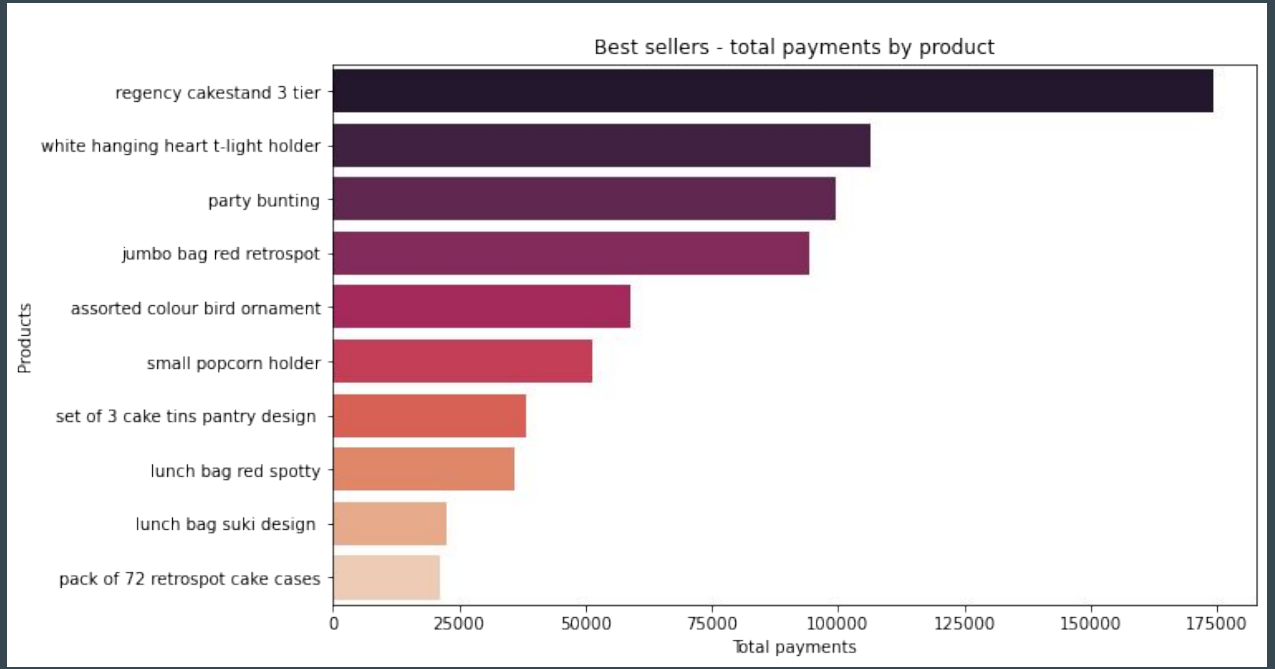
Recency

They were all purchased on the most recent date

The 4 most gainful best sellers are the most gainful overall

Findings

- The 5th is also among the 10 most gainful overall
- The most gainful product gained substantially more revenue than the others (8 times more than the 10th)



Their purchases varied similarly to the general time trend

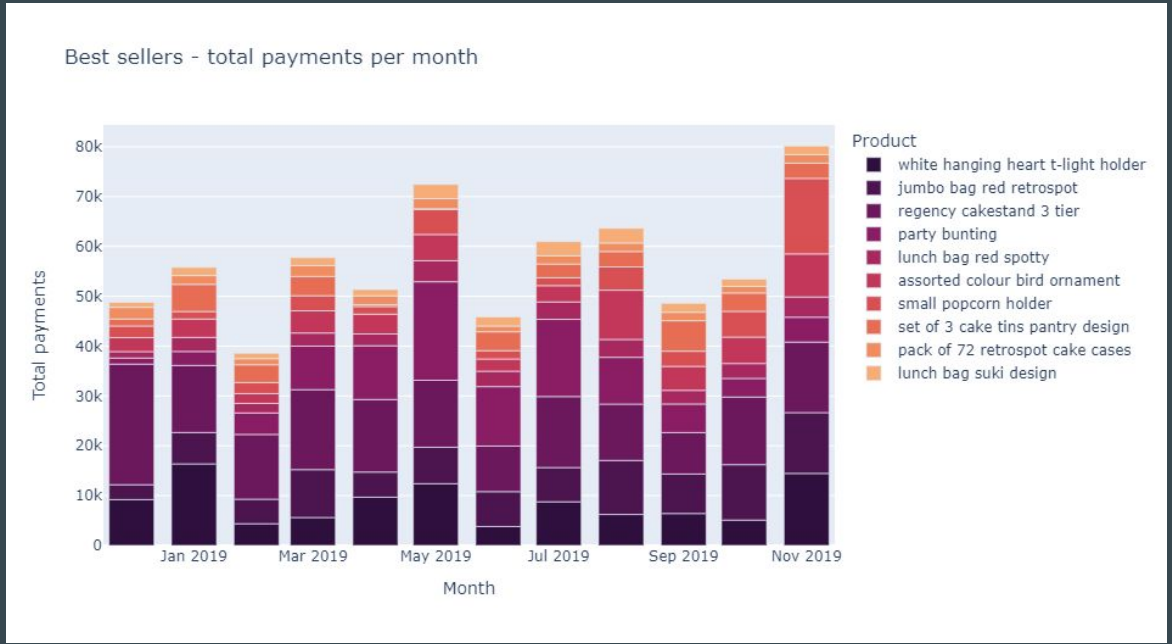
Findings

- The share of the different products was relatively similar
- The best seller product was leading on: December 2018, January 2019, November 2019

Their revenue was less affected by seasonality

Findings

- July-August did better than September-October
- The share of the different products was more diverse. Most of the revenue belonged to the top 4 products
- The most gainful product took the lead most of the time



The most gainful product is the most expensive best seller

Findings

- The most gainful product was substantially more expensive than the others (17 times more than the 10th)
- It sold few units compared to the others
- The best seller product is situated in the middle

* Average quantity & price of each product

The best sellers exceeded other products in quantity & price

- Their advantage was clearer in terms of quantity

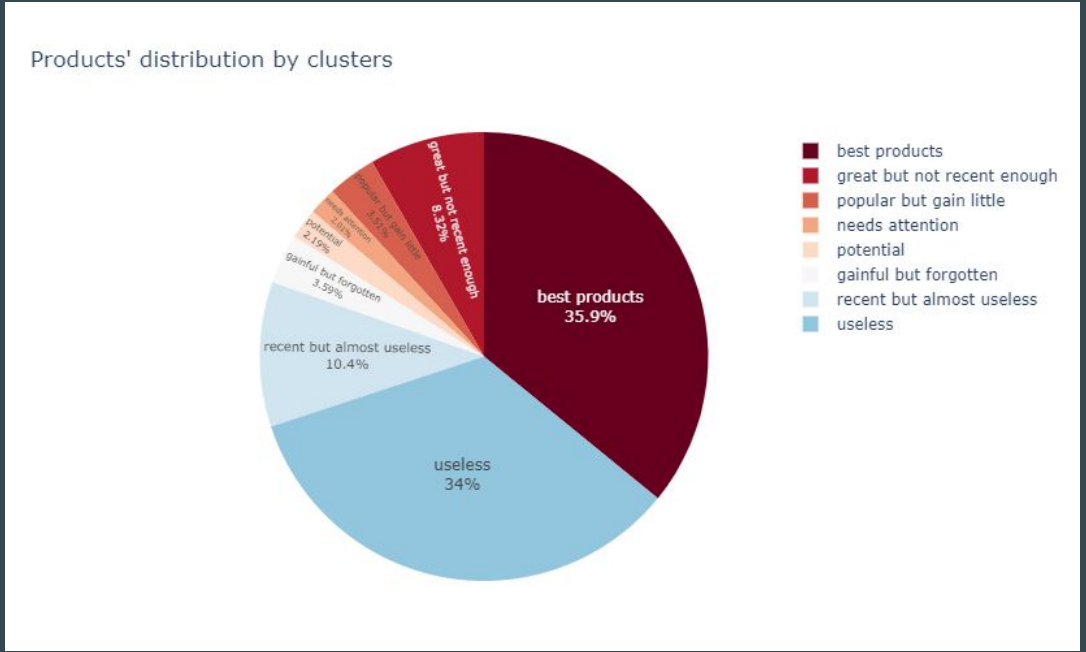
	Quantity		Price	
	Mean	Median	Mean	Median
Best sellers	17 units	6 units	4.2 $	2.5 $
Other products	10 units	4 units	3.2 $	2.1 $

What can we learn from segmenting products?

At least 34% of the products are unsuccessful

Findings

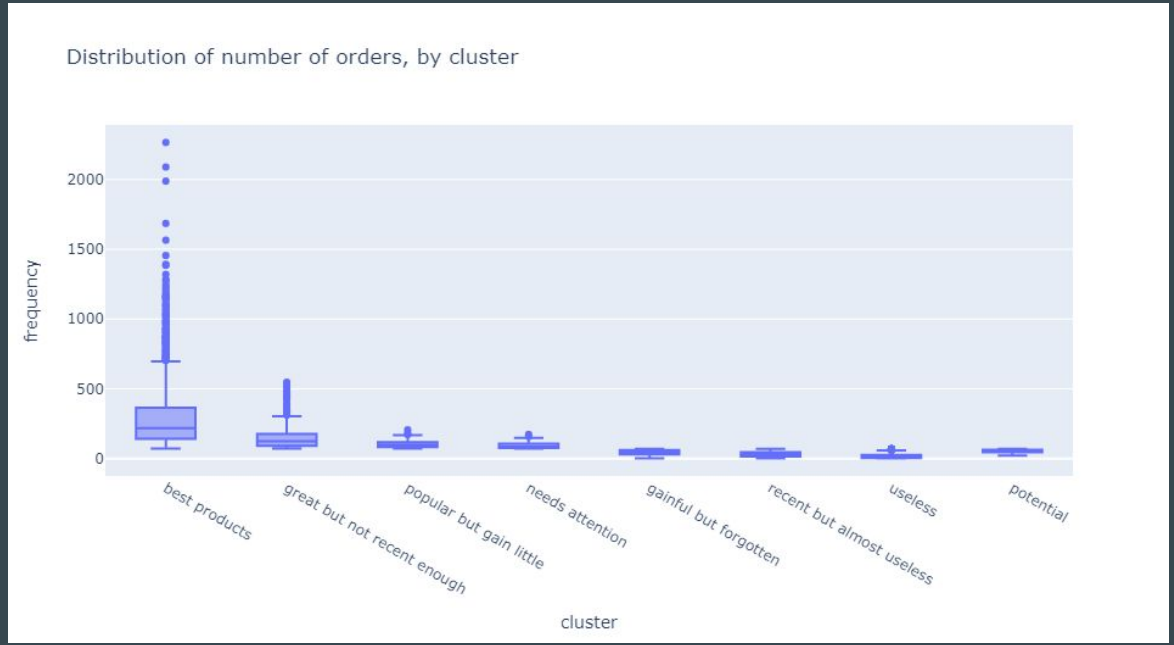
- 34% can be clustered as "useless" (with low purchases, low revenue, and long time since purchase)
- 30% more require different levels of attention
- 36% are "best products", that excel both in purchases and in revenue



The “best products” cluster had the highest purchases

Findings

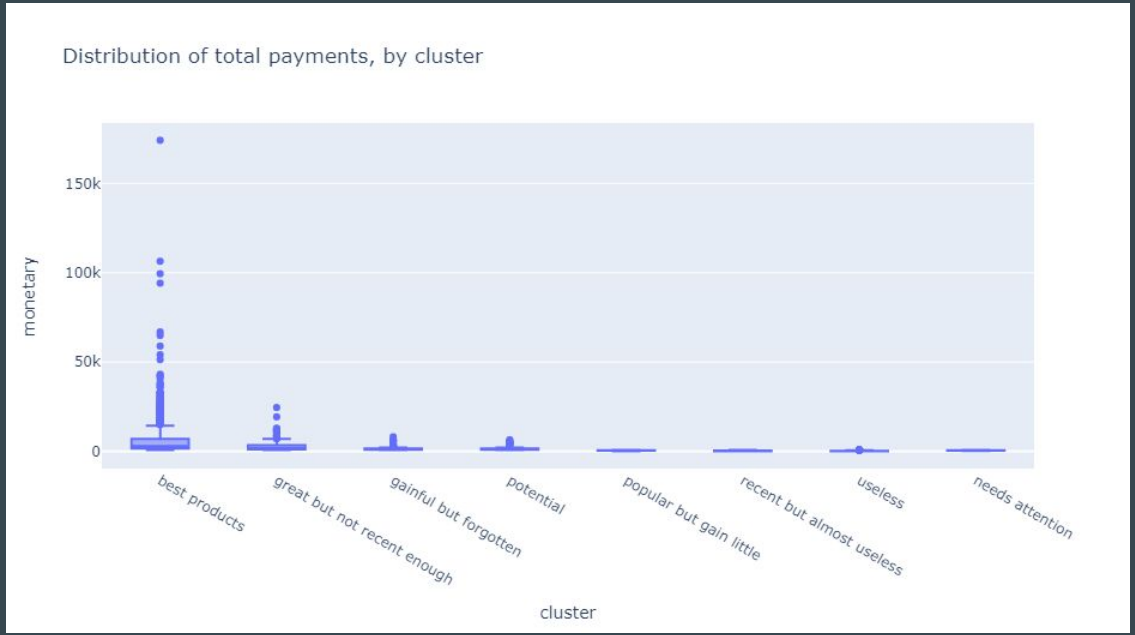
- Its median is the highest and it has the highest outliers
- It’s also the most diverse
- "Best products" vs. "useless": 297 purchases vs. 17, on average



The “best products” cluster gained the highest revenue

Findings

- "Best products" vs. "useless": 6,123$ vs. 161$, on average
- The distribution of the revenue is less diverse than that of the purchases



The purchases of the “best products” varied by the season

Findings

- Their time trend was similar to the general trend
- The trend of the following cluster is quite similar, but its purchases did not rise in November as much
- The rest of the clusters showed fewer fluctuations over the year

The revenue of the “best products” varied by the season

Findings

- Their time trend was similar to the general trend
- The rest of the clusters showed almost no fluctuations over the year
- The revenue of the “best products” was substantially higher than the others’

The quantity played a more prominent role than the price

Findings

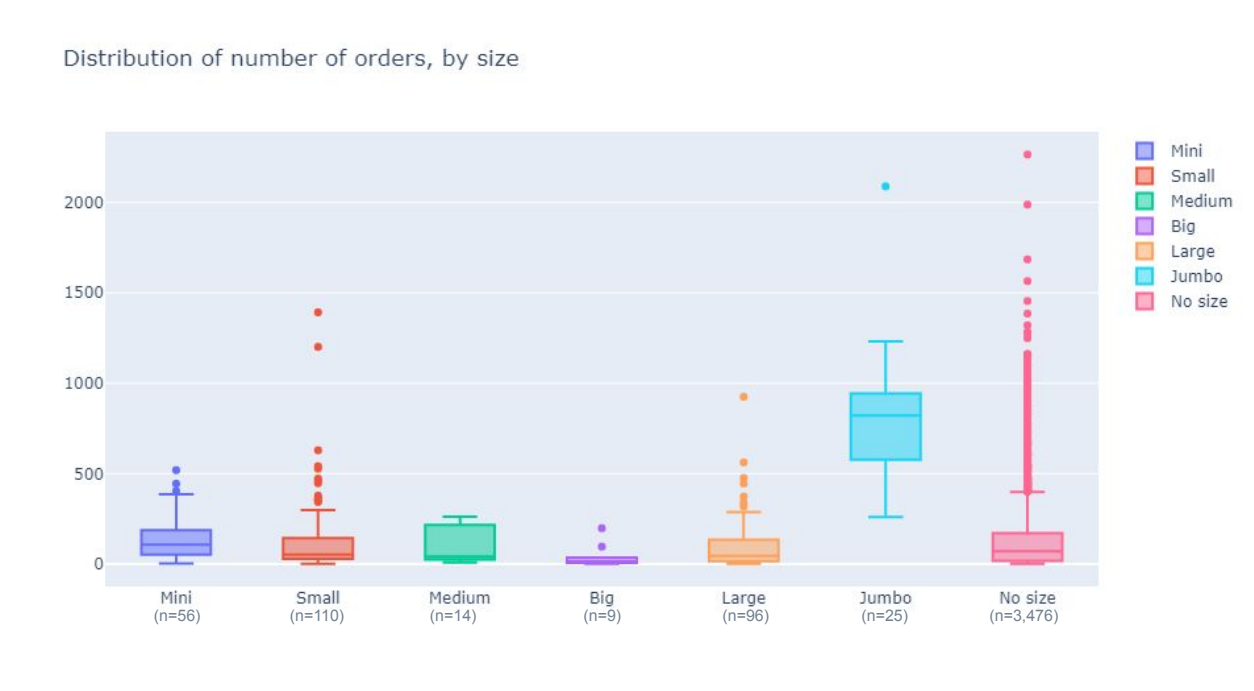
- The “best products” sold an average of 10 units, while the “useless” products sold 6.8 units
- However, the average price of the “best products” was lower (3.3$ vs. 3.9$)

Does the size matter?

“Jumbo” products had higher purchases than other sizes

Findings

- "Mini" is the second in line
- 'Medium' and 'big' have too few products, and their distribution is too different to merge
- We shall focus on “jumbo”, “mini” and “others”



“Jumbo” products gained higher revenue than “mini” & “others”

Findings

- “Mini” products gained higher median revenue than the “others”, but the “others” include extremely gainful outliers

“Jumbo” products excelled in purchases, revenue & quantity

Purchases

- They accounted for 4% of purchases
- Their purchases were significantly more likely to be higher than those of “mini” and “other” products

(Mann–Whitney, $p < 0.01$, two-tailed)

Revenue

- They accounted for 6% of the revenue
- Their revenue was significantly more likely to be higher than that of “mini” and “other” products

(Mann–Whitney, $p < 0.01$, two-tailed)

Quantity

- They sold more units than others (14 units vs. 10, on average)
- They were cheaper than others (2.5 $ vs. 3.3 $, on average)

Just so you know...

The data

The transaction history of an online store that sells household goods

- **25,900 invoices**
- **Timeline:** 29.11.2018 - 7.12.2019
  (no data on 22.12.2018-1.1.2019 and on 20-23.4.2019, most likely because of the holidays)

Data contains:

- Invoice identifier and date
- Item identifier, name, quantity and price per unit
- Customer ID

Check out the dashboard