E-Commerce EDA & Clustering

Quarter of Order Date

Discount category
- 0 Discount
- > 0.20 Discount
- ≤ 0.20 Discount

Company makes a profit on full-price items, a small profit on items reduced by 20% or less, and a loss on big-discount items.

E-Commerce EDA & Clustering

Quarter of Order Date

Discount category
- 0 Discount
- > 0.20 Discount
- ≤ 0.20 Discount

Discounted items consistently make up around 50% of sales, full-price sales the other 50%. Within the discounted sales, the majority are large discounts of over 20% off.

E-Commerce EDA & Clustering

Cluster labels

Discount category
- 0 Discount
- > 0.20 Discount
- ≤ 0.20 Discount

Our customer clusters all have very similar preferences when it comes to discount buying behaviour.

E-Commerce EDA & Clustering

Quarter of Order Date

Cluster labels
- Outliers (125)
- Frequent buyers (250)
- Lapsed (103)
- New customers (315)

Frequent buyer cluster makes up a large portion of sales. Lapsed customers represent a missed opportunity for sales. How to win them back?

E-Commerce EDA & Clustering

Sub Category

Cluster labels
- Outliers (125)
- Frequent buyers (250)
- Lapsed (103)
- New customers (315)

The outlier customer cluster has made big purchases in certain subcategories. Otherwise, hard to discern difference in subcategory preference of clusters. Certain subcategories are unprofitable.

E-Commerce EDA & Clustering

Sub Category

Avg. Profit margin

Cluster labels
- Outliers (125)
- Frequent buyers (250)
- Lapsed (103)
- New customers (315)

Mostly even distribution between customer groups. Profit made on machines bought by lapsed customers; loss made on machines bought by other custome groups. Lapsed customers got a lot of bargain bookcases.

E-Commerce EDA & Clustering

Key Takeaways:

So far, we haven't identified significant differences in buying behaviour of different customer clusters.

Recommondation: Try another clustering approach trained on different features, then compare again with overall trends.

Aim: Identify marketing measures that will be most effective on each customer cluster.