Project E-Commerce: Product Range Analysis

Project Decomposition by Yasmin Madjitey

Project Goal:

The main goal here is to analyze the store's product range with regards to profitability and customer demand. That is, a careful analysis to provide the basis for targeted adjustments to boost sales, hence profits.

Scope Analysis: Task

- Carry out exploratory data analysis:

1) Data Preprocessing. That is, general data cleaning like casting columns into the right data type, check for duplicates, missing values, etc.

2) Analyze data distribution for individual variables, correlation between variables and product association rules.

- Analyze the product range:

The main task here encompasses the analysis of store orders and or products. For example, which products are ordered more frequently than others, assorted products, and which unrelated store items are ordered as a consequence of ordering another item.

- Formulate and test statistical hypotheses:

1) Problem statement: How stable is the customers' demand?

Hypotheses: can it be explained by demand variability or seasonality?

Answering this problem probably by way of looking at the distribution of sales for an item or better still, knowing if demand for products follows a normal distribution, can go a long way to help the store prepare for peak sales. As we can imagine, it will be way more challenging to handle peak sales than a uniform distribution throughout a proposed time period. This way, it will be easier to understand which products will bring planning and distribution challenges so as to plan ahead of time.

2) Problem statement: What items are driving most sales?

Hypotheses: From the problem statement above, popularity comes to play, as well as, categorization of items based on fast movers, slow movers, and very slow movers (RFM metrics can be used here for categorization). From this categorization, I can plot demand variability and the distribution of sales for each group.

This problem statement is with regards to products in high demand as well as products that drive revenue leading to a follow-up question: why are these items frequently ordered? could it be because they cost lower than other items? (Here, I will be looking at affordability as against profitability. I can plot the most popular items or fast selling items against price to understand if price affects sales.)

3) Problem statement: Does price affect sellability?

4) Problem statement: What products are most often sold together?

Hypotheses: This raises the topic of association rules. With this knowledge, the store can drive sales by probably recommending associated items to customers or apply discounts to the associated items to drive sales.

Conclusion

I strongly believe that the above mentioned problem statements will provide some insights to curb challenges the store will face in terms of planning, distribution and management of the goods flow to meet demand.

References

During my research on the topic, the following websites enlightened me on some of the questions a data analyst should aim to answer when conducting product range analysis.

- https://towardsdatascience.com/product-segmentation-for-retail-with-python-c85cc0930f9a
- https://towardsdatascience.com/shop-order-analysis-in-python-ff13615404e0
- https://medium.com/swlh/product-sales-analysis-using-python-863b29026957
- https://sunscrapers.com/blog/sales-data-science-a-step-by-step-guide-to-competitor-analysis-using-python/