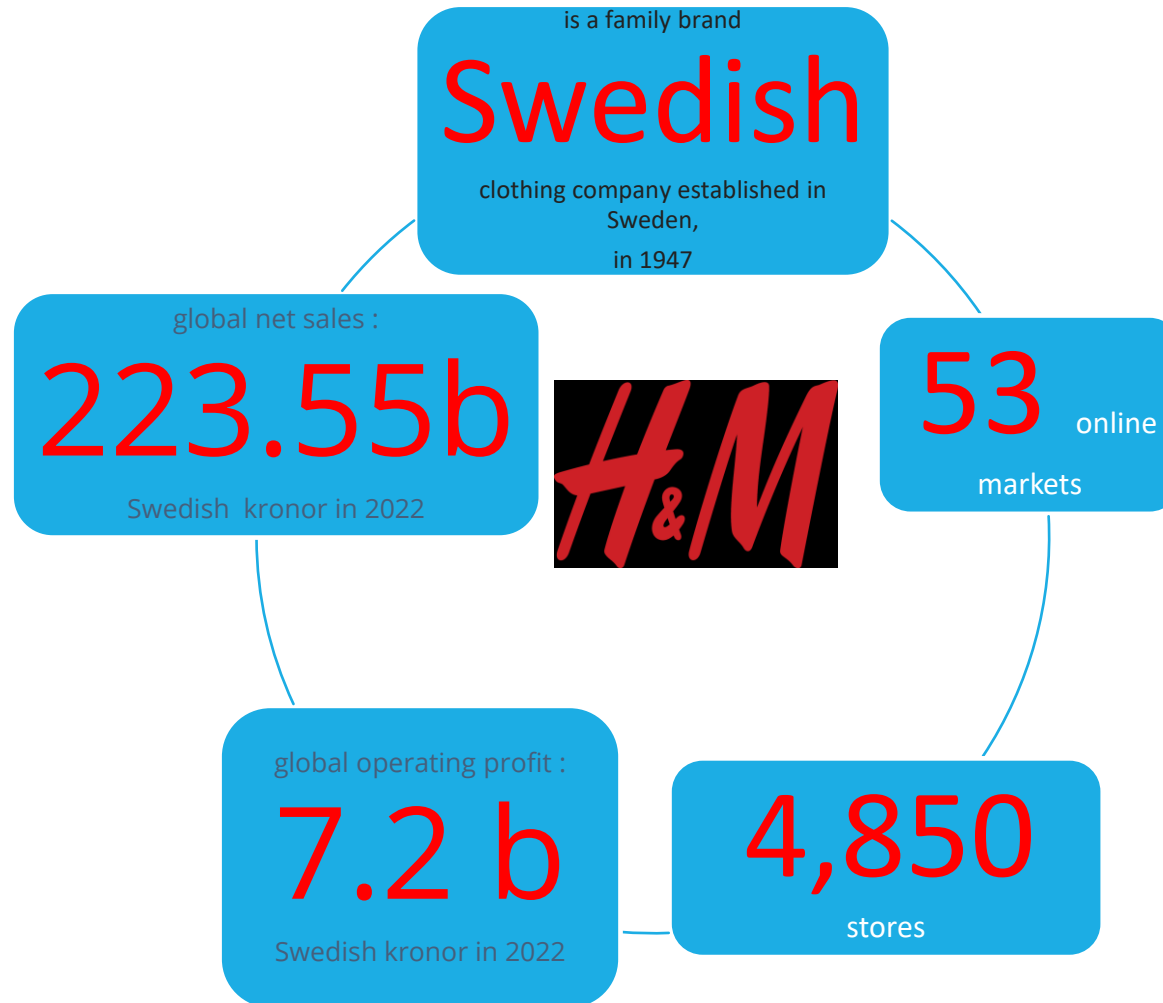

H & M RECOMMENDER SYSTEM

CAPSTONE PROJECT

BY ALICE FROM DSI39



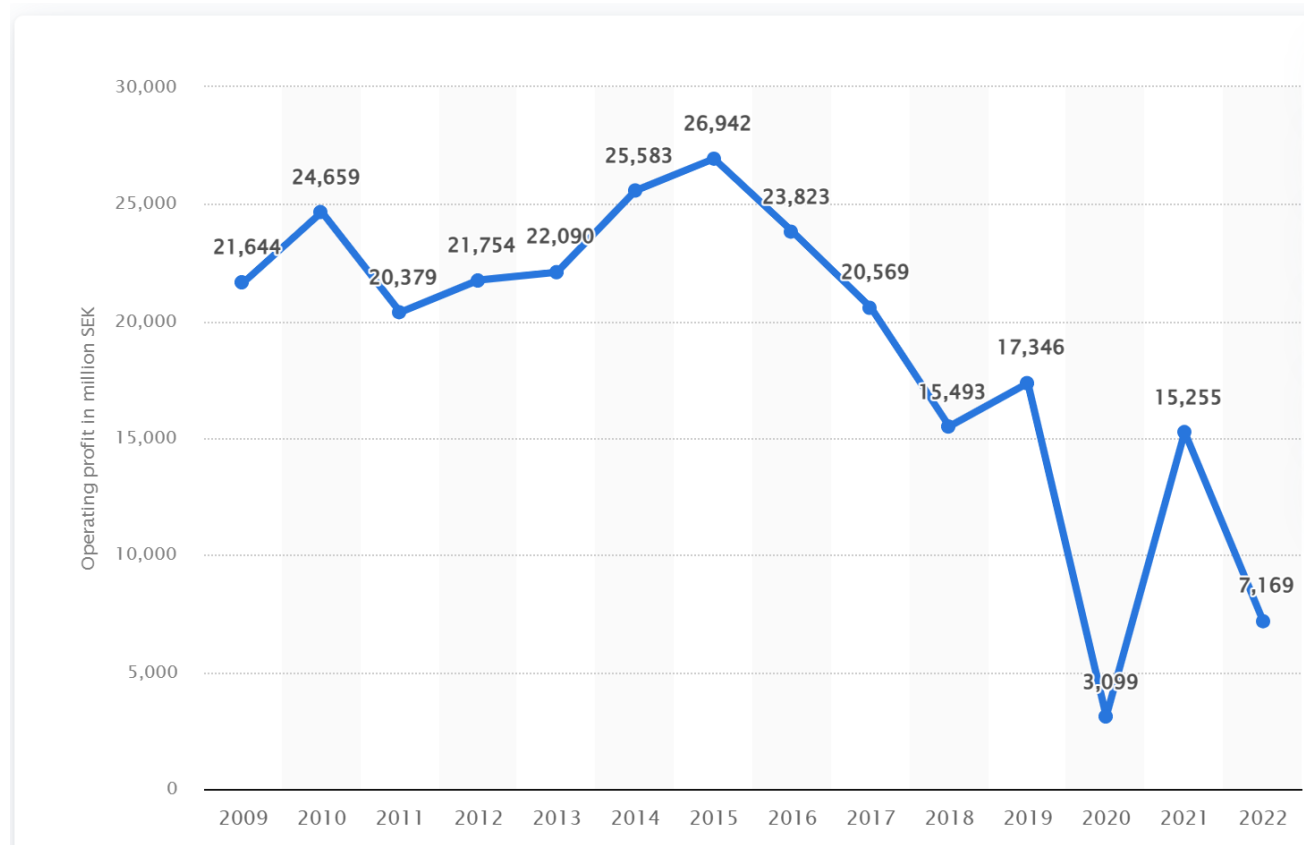
SIGNIFICANCE OF PROJECT



BASICS SOLD BY H&M



OPERATING PROFIT OF THE H&M GROUP WORLDWIDE



Reference: statista

PROBLEM STATEMENT

- H&M customers have too many products to browse online.
- While some customers are good at searching for what they want, others may prefer not to do so.
- In the end, they might not make a purchase.
- To help customers make quick purchases, product recommendations become important.
- This project aims to make product recommendations to H&M customers in order to help them decide which products to purchase.

IMPACT OF PRODUCT RECOMMENDATION

- The recommendation engine market register a CAGR of **37.46%** during the period of 2021-2026
- Shoppers that clicked on recommendations are **4.5x** more likely to add the items to their cart and **4.5x** more likely to complete the purchase. (Salesforce)
- Shoppers who are engaged with a recommended product had a **70%** higher conversion rate during that session. (Monetate)
- product recommendations make up for **24%** of orders and **26%** of revenue. (Salesforce Research)
- Personalized product recommendations are estimated to account for more than **35%** of purchases on Amazon

Reference: The Power of Product Recommendations: 30 Must-Know Statistics for 2022
(clerk.io)

META DATA

4 files are provided for the competition:

Customer data	Age, postal code, club membership status, fashion news frequency
Articles data	24 description of products in code and name e.g. article 0108775015 is a strap top, vest top, garment upper body, solid, black, dark, jersey basic, ladies wear, womens everyday basics, Jersey top with narrow shoulder straps.
Transactions data	Transaction date, customer id, article id , price, sales channel
images	115 000 images of products, include clothes, accessories, shoes and bags



DATA CLEANING

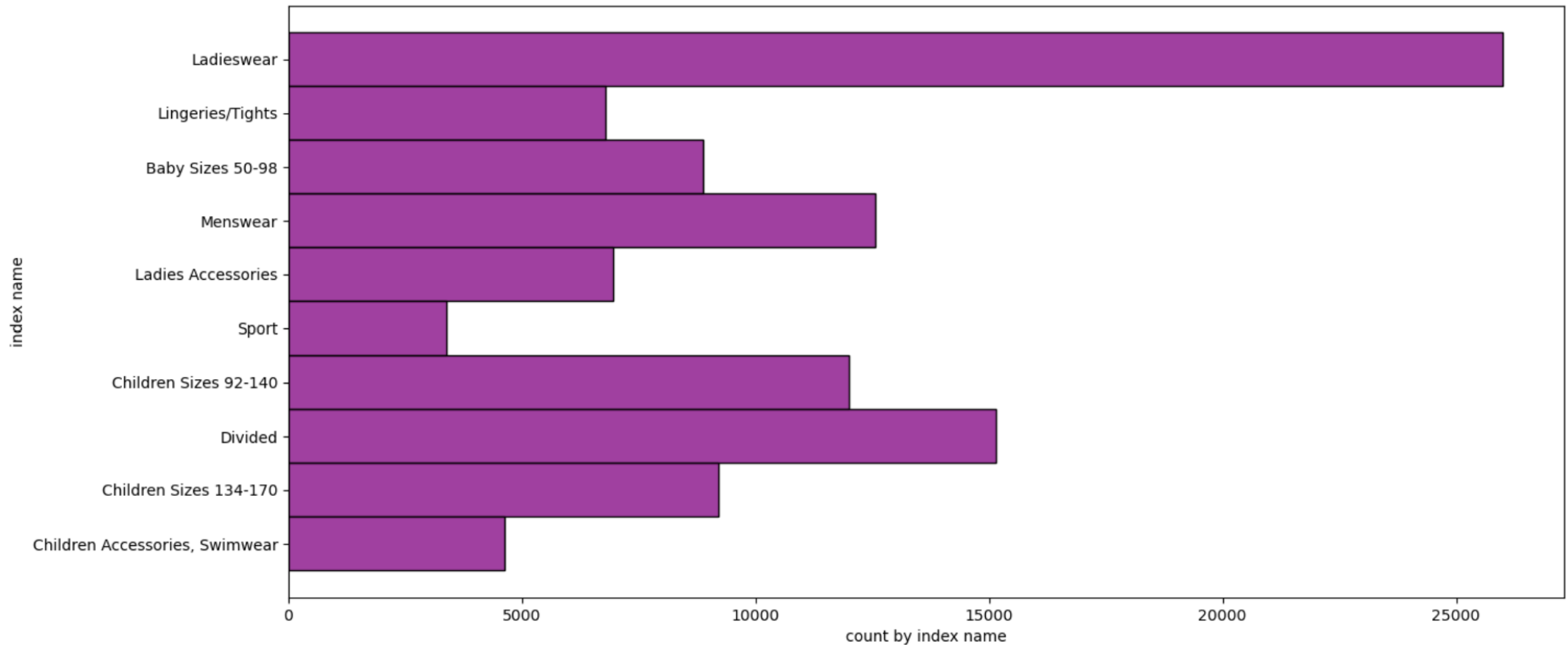
- Most data are clean except for null values in:
 - detail_desc
 - FN
 - Active
 - club_member_status
 - fashion_news_frequency
 - age
- rows with null values are dropped

EDA - COUNT OF PRODUCTS SOLD BY INDEX NAME

No. 1 : Ladieswear

No.2: Divided

No.3 : Menswear

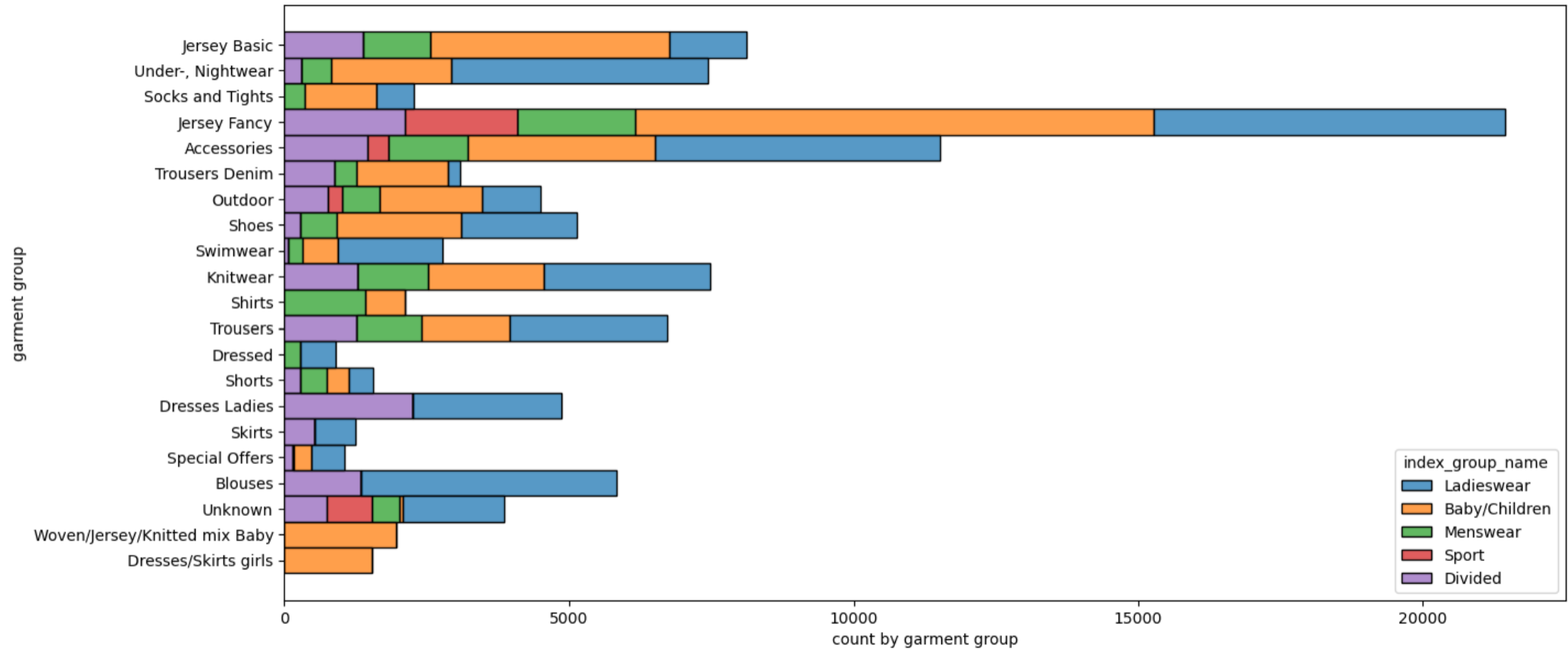


EDA - COUNT OF PRODUCTS SOLD BY GARMENT GROUP

No. 1 : Jersey Fancy

No.2: Accessories

No.3 : Jersey Basic

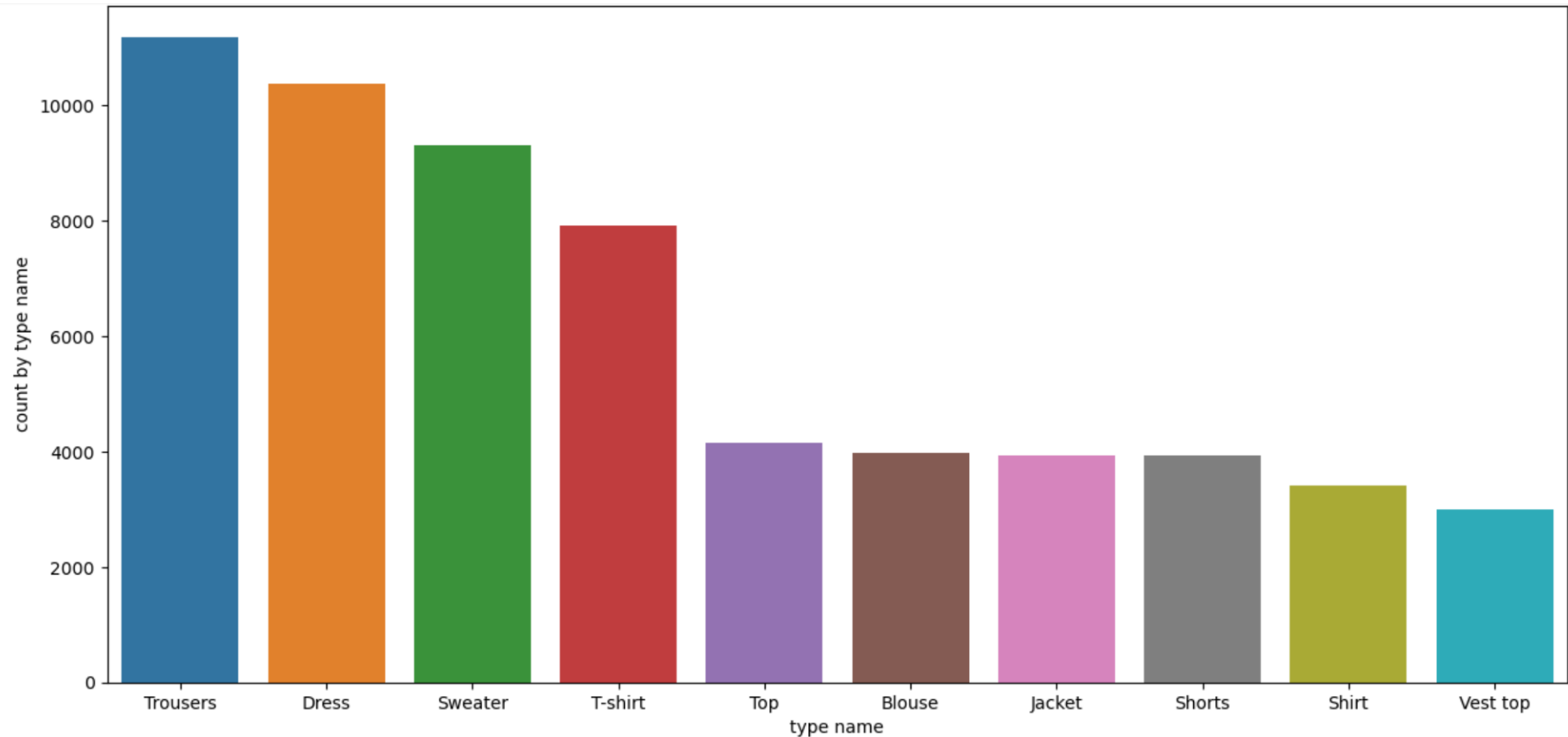


EDA - COUNT OF PRODUCTS SOLD BY TYPE NAME

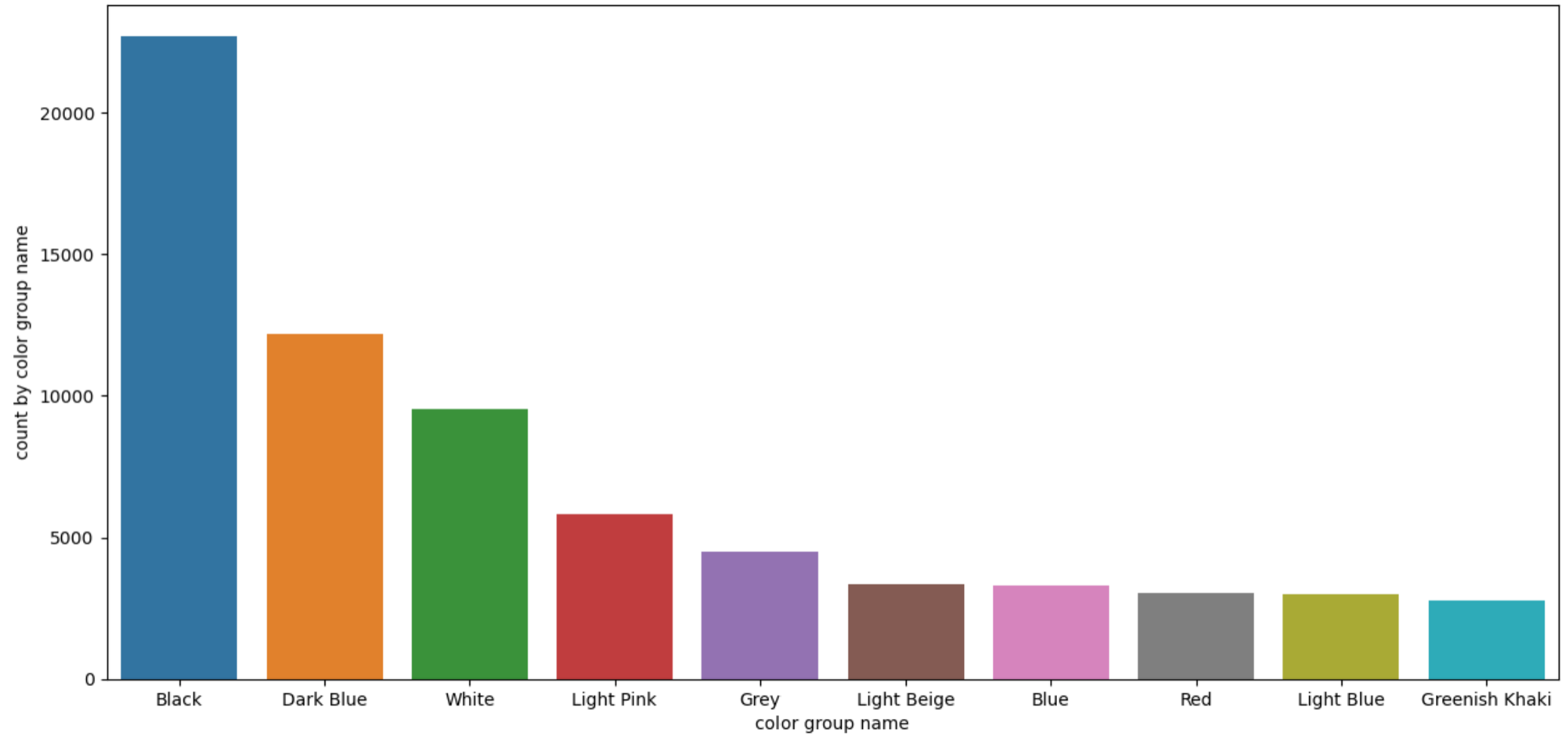
No. 1 : Trousers

No.2: Dress

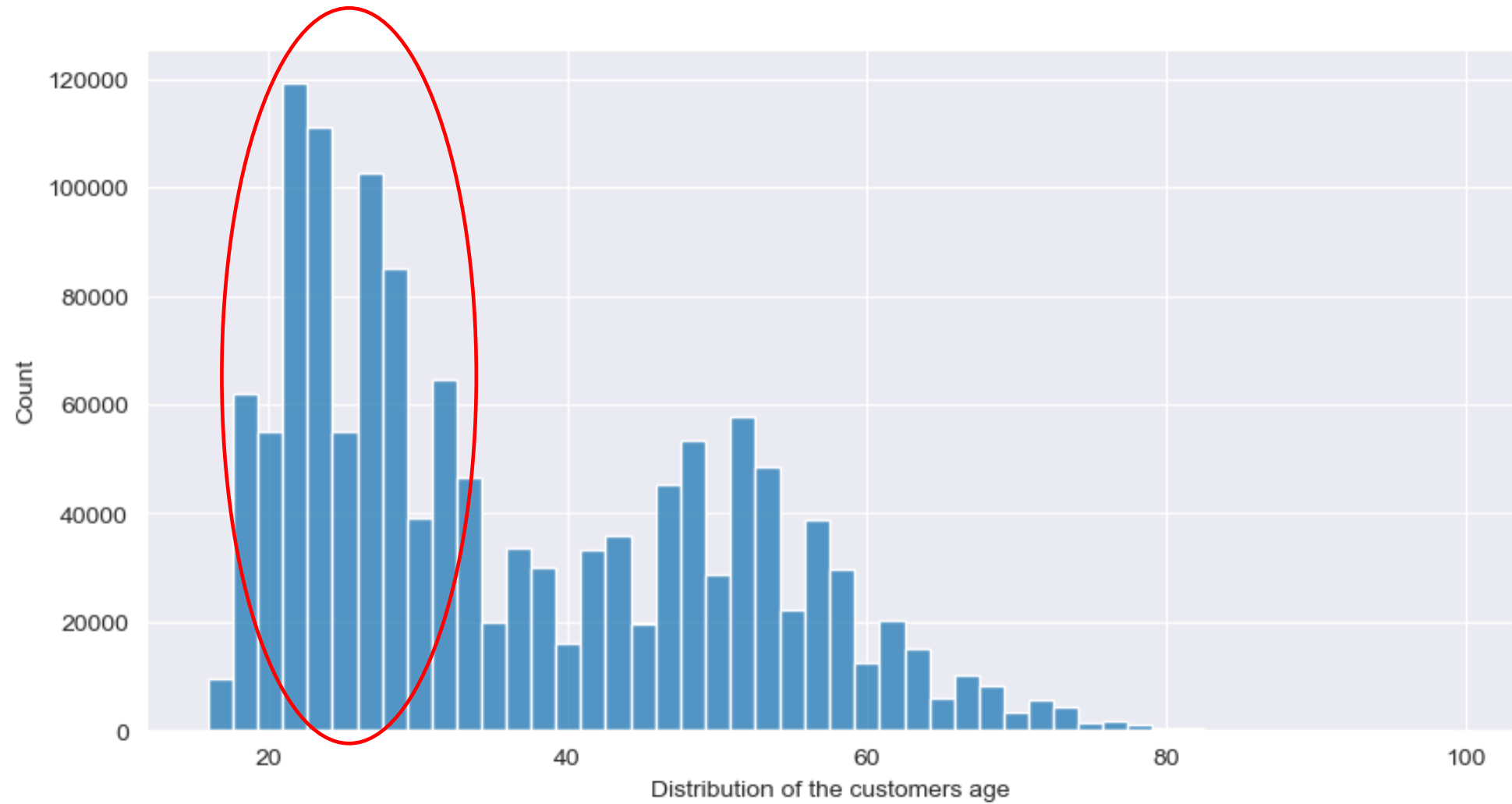
No.3 : Sweater



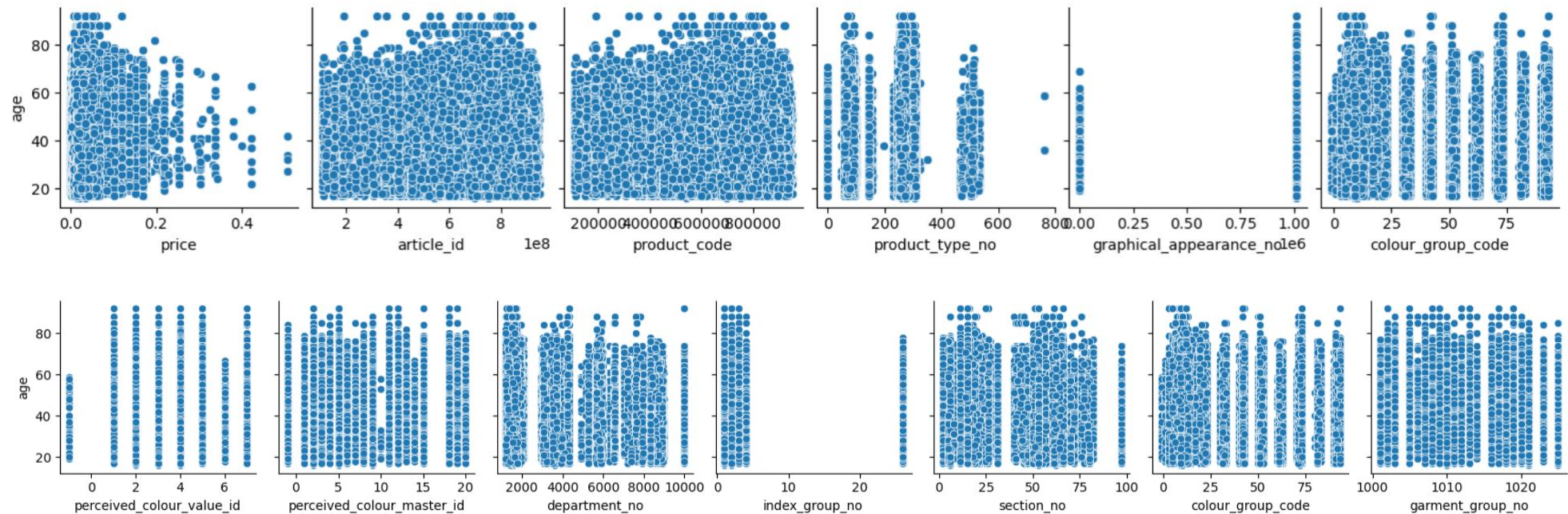
EDA – COLOUR BY POPULARITY



EDA – CUSTOMERS AGE



EDA – AGE AND OTHER VARIABLES



Age and price seems to have a correlation. Most expensive item is bought by those aged between 30 to 40.
Low priced items are bought by people of all ages.
As price increases, those outside the age range of 20 to 75 do not buy them.

RECOMMENDER SYSTEM BY CONTENT-BASED FILTERING

2 Types of Recommender System:

1. Content-Based

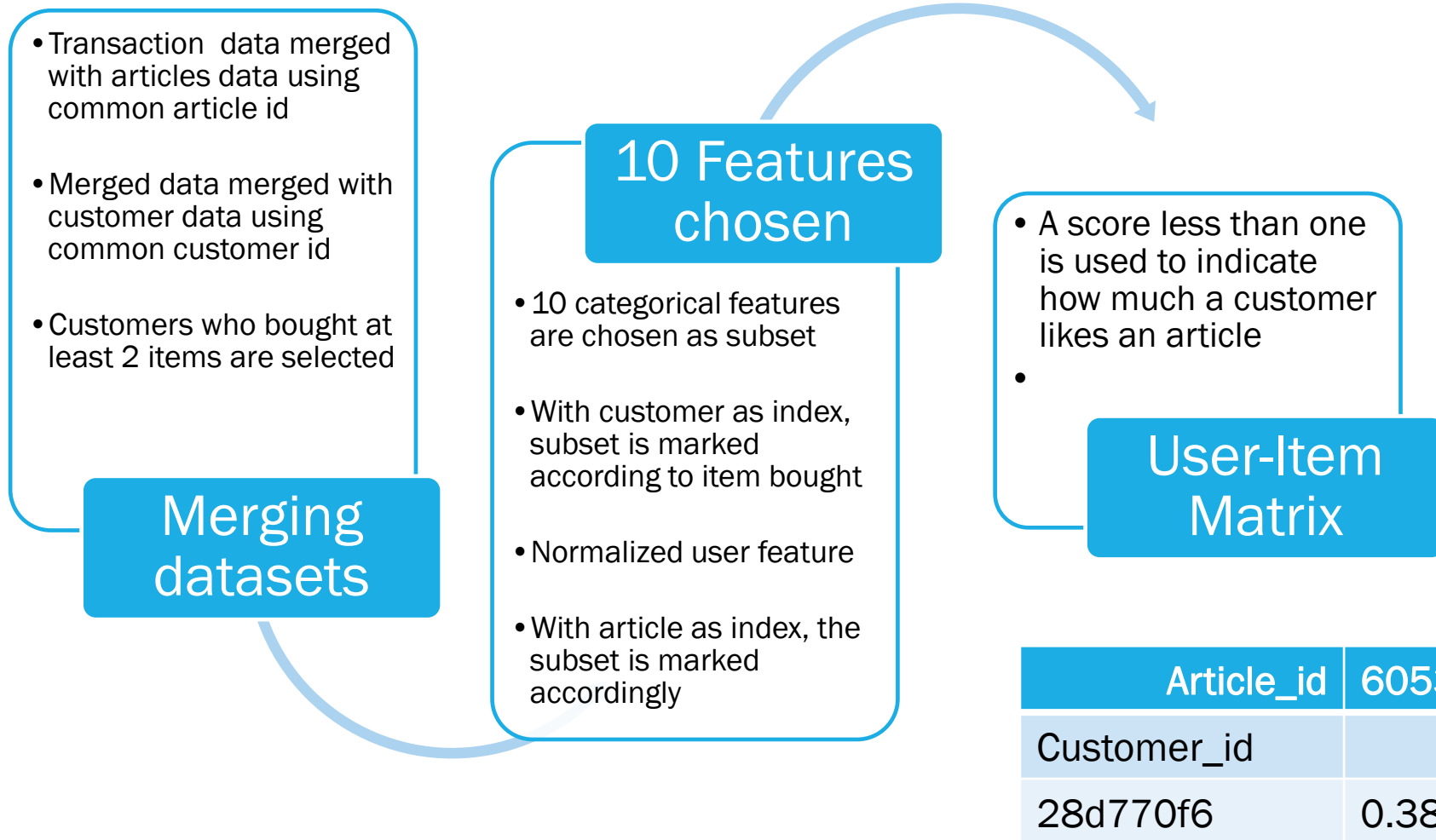
- **Implicit Rating**, suggests users preference indirectly, such as page views, clicks, purchase records

2. Collaborative

- **Explicit Rating**, is a rate given by a user to an item on a sliding scale, like 5 stars for Titanic.

Reference: Introduction to Recommender System , Towards Data Science

HOW THE RECOMMENDER SYSTEM IS BUILT





CHALLENGES

- Very large data set is given so not all datasets are used
- only 10000 customers data are used
- as categorical features are used, large memory is needed to build the matrix

CONCLUSION

- Recommender System gives recommendation to customers based on their past purchases
- It cuts down the time needed for customers to find products which might suit them
- Helps H&M company to increase her profit when customers can find what they want through the online recommendation



STREAMLIT DEMO

Streamlit Link