PERSONALISED RECOMMENDATIONS FOR H&M CUSTOMERS

Helen Gaskell

PROBLEM STATEMENT

- ➔ How can we make it easier for customers to find exactly what they’re looking for?
- ➔ How can we recommend items that a customer might be interested in?

WHY IS THIS USEFUL?

- ➔ Quicker and more efficient buying process
- ➔ Reduce returns and transport emissions
- ➔ Increase revenue by increasing the number of purchases per customer

DATASETS: PROVIDED BY H&M

→ TARGET VARIABLE: article_id

CUSTOMERS: 3 million

TRANSACTIONS: 30 million

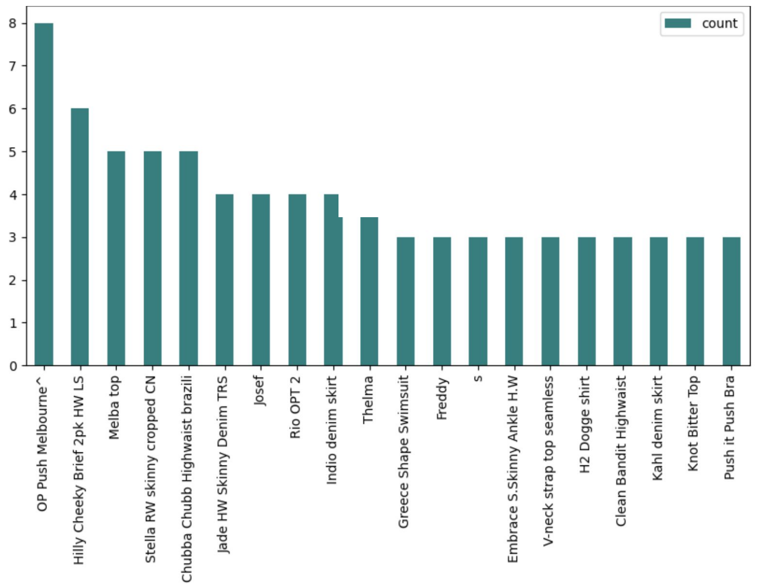
ARTICLES: 100,000

EXPLORATION & ANALYSIS

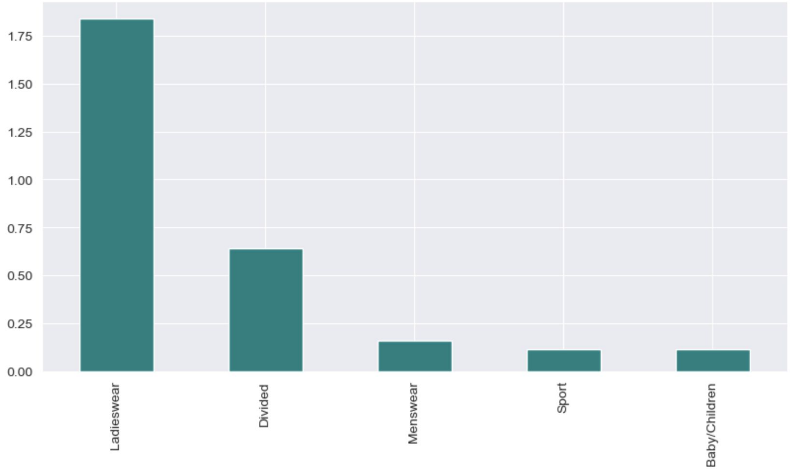
Top customer bought 1346 items →

Repurchased top product 8 times →

H&M'S TOP CUSTOMER PURCHASES



TRANSACTIONS BY DEPARTMENT (MILLIONS)



← Ladieswear counts for over 70% of total transactions

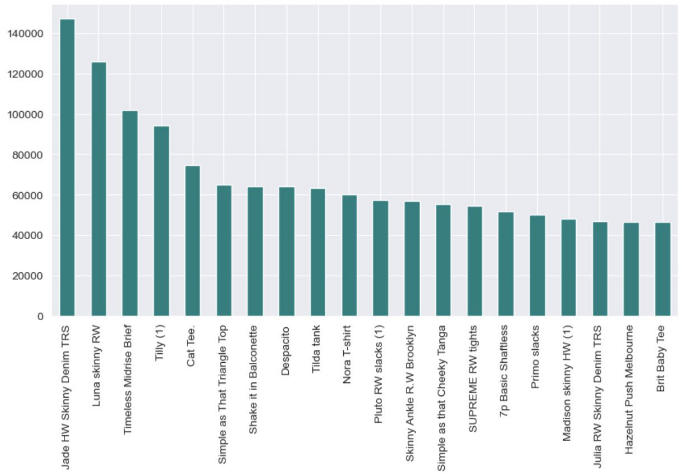
← 8 times more ladieswear purchased than menswear

EXPLORATION & ANALYSIS

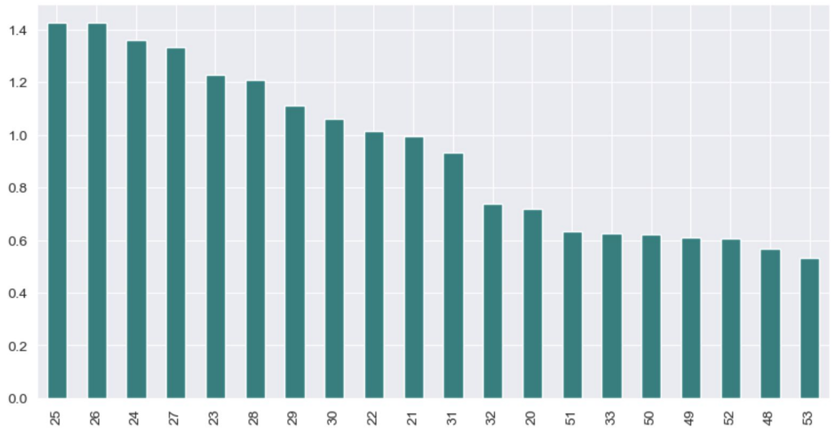
Skinny jeans are H&M’s most popular product →

Most popular product was a repeated purchase of top customer →

H&M’S TOP PRODUCTS (TRANSACTIONS)



TRANSACTIONS BY AGE (MILLIONS)



← Ages 24-26 make the most purchases

MODELING

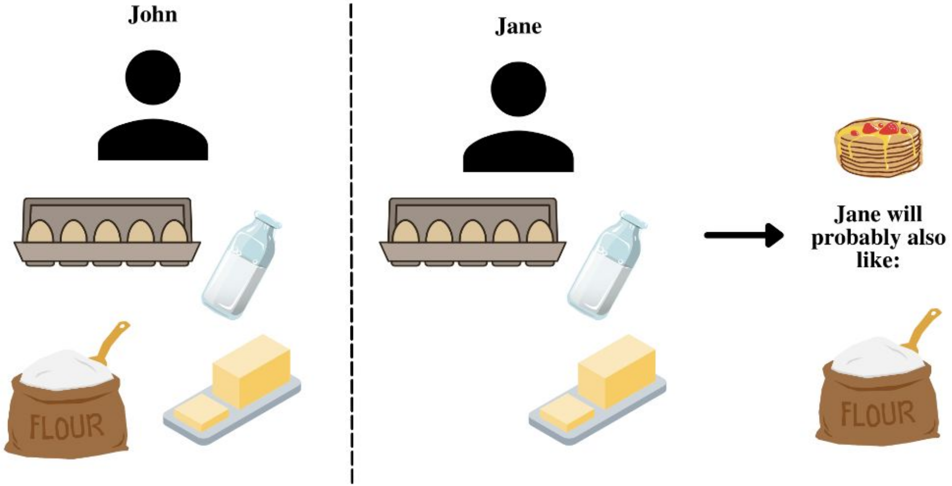
BASELINE MODEL:

- ➔ MOST POPULAR ITEMS
- ➔ RECOMMEND MOST POPULAR ITEM WITHIN A CUSTOMER'S MOST FREQUENTLY PURCHASED PRODUCT TYPE

ADVANCED MODELING:

- ➔ DECISION TREE
- ➔ MARKET BASKET ANALYSIS:
  - ◆ COSINE SIMILARITY
  - ◆ PEARSON SIMILARITY

COLLABORATIVE FILTERING



3 INPUT TABLES:

- ➔ ARTICLE PURCHASE COUNT
- ➔ PURCHASE DUMMY
- ➔ NORMALIZED PURCHASE COUNTS

RESULTS

MOST ACCURATE MODEL?

➔ COSINE SIMILARITY

WHICH TABLE?

➔ PURCHASE DUMMY

ACCURACY METRICS?

➔ RMSE & PRECISION AND RECALL