

# ManyGiftsUK Recommender System

Ana Luísa Mestre    Beatriz Pereira  
Mariana Domingues    Nadine Aldesouky



# Business Objective

Build a recommender system that can facilitate user choices by:

- suggesting items that users would like
- solving the cold start problem



# DATASET

- Retail csv file containing 8 variables and 541909 records
- Records were taken between 01/12/2010 and 09/12/2011
- The dataset includes:
  - 25900 valid transactions
  - 4070 unique items
  - 4372 customers
  - 38 different countries.

- From the total of 25900 transactions, 3836 were **canceled** (14.81%)

- 10624 instances have **negative quantities**



9288  
Cancellations

1336  
Non-Cancellations

- 19265 **deduplicated** instances by InvoiceNo and StockCode



Aggregate instances  
by summing **Quantity**

- Transactions come from **38 different countries** of which:

- 24 are European
- 8 are Asian
- 3 are American
- 1 is African
- 1 is Oceanic
- 13 are unspecified

- From the total of 25900 transactions, 23494 come from the **UK** (90.71%)

# INSIGHTS



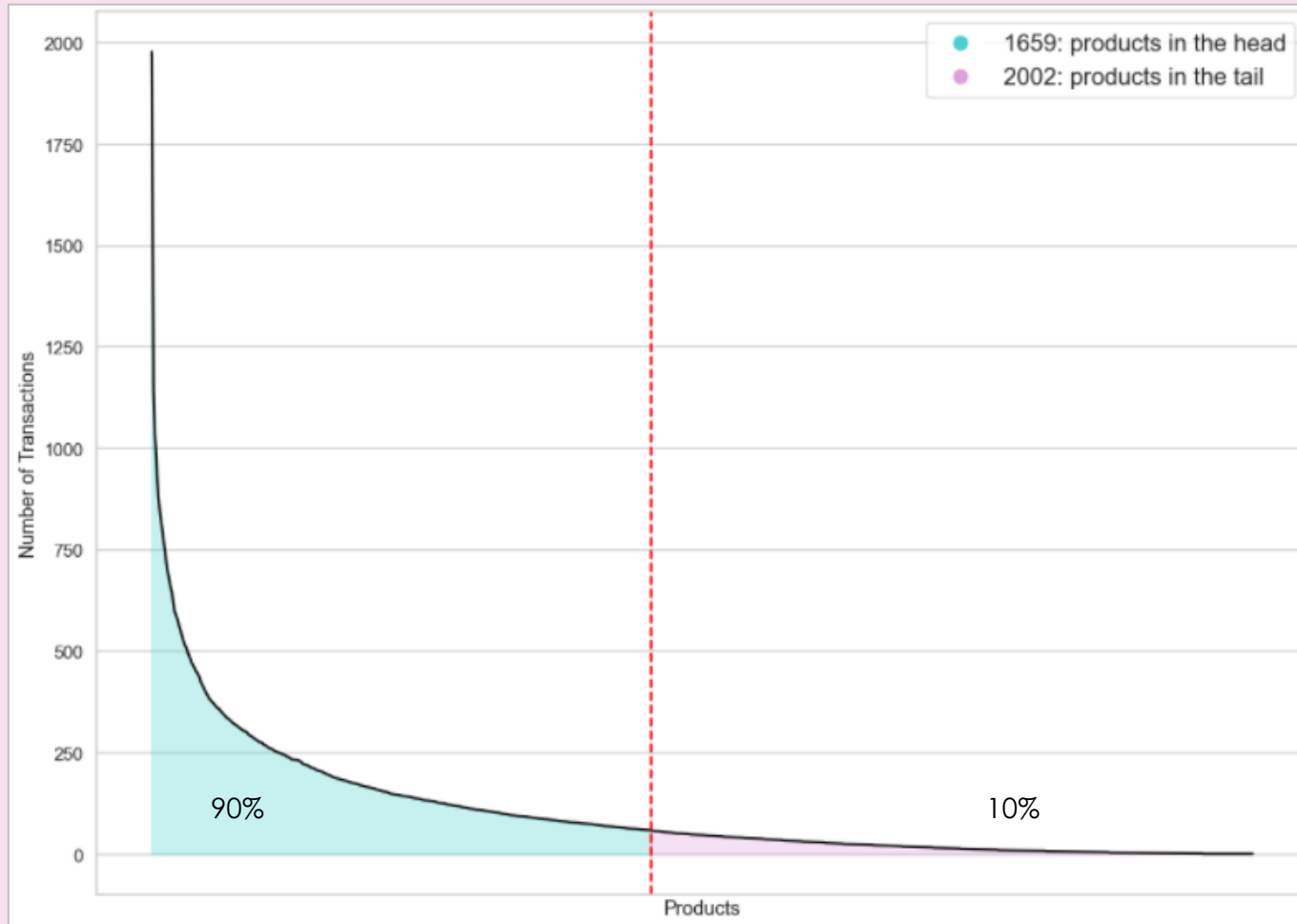
# INSIGHTS

- Stock Code: 5-Digit Code
- 10 products with a shorter code
- Drop all except PADS

## Doubtful Products

Stock Code	Description	Frequency (Number of observations)	Variety of Prices
POST	POSTAGE	1256	114
DOT	DOTCOM POSTAGE	710	687
M	Manual	571	260
C2	CARRIAGE	144	6
D	Discount	77	75
S	SAMPLES	63	59
CRUK	CRUK Commission	16	16
PADS	PADS TO MATCH ALL CUSHIONS	4	2
m	Manual	1	1
B	Adjust bad debt	1	1

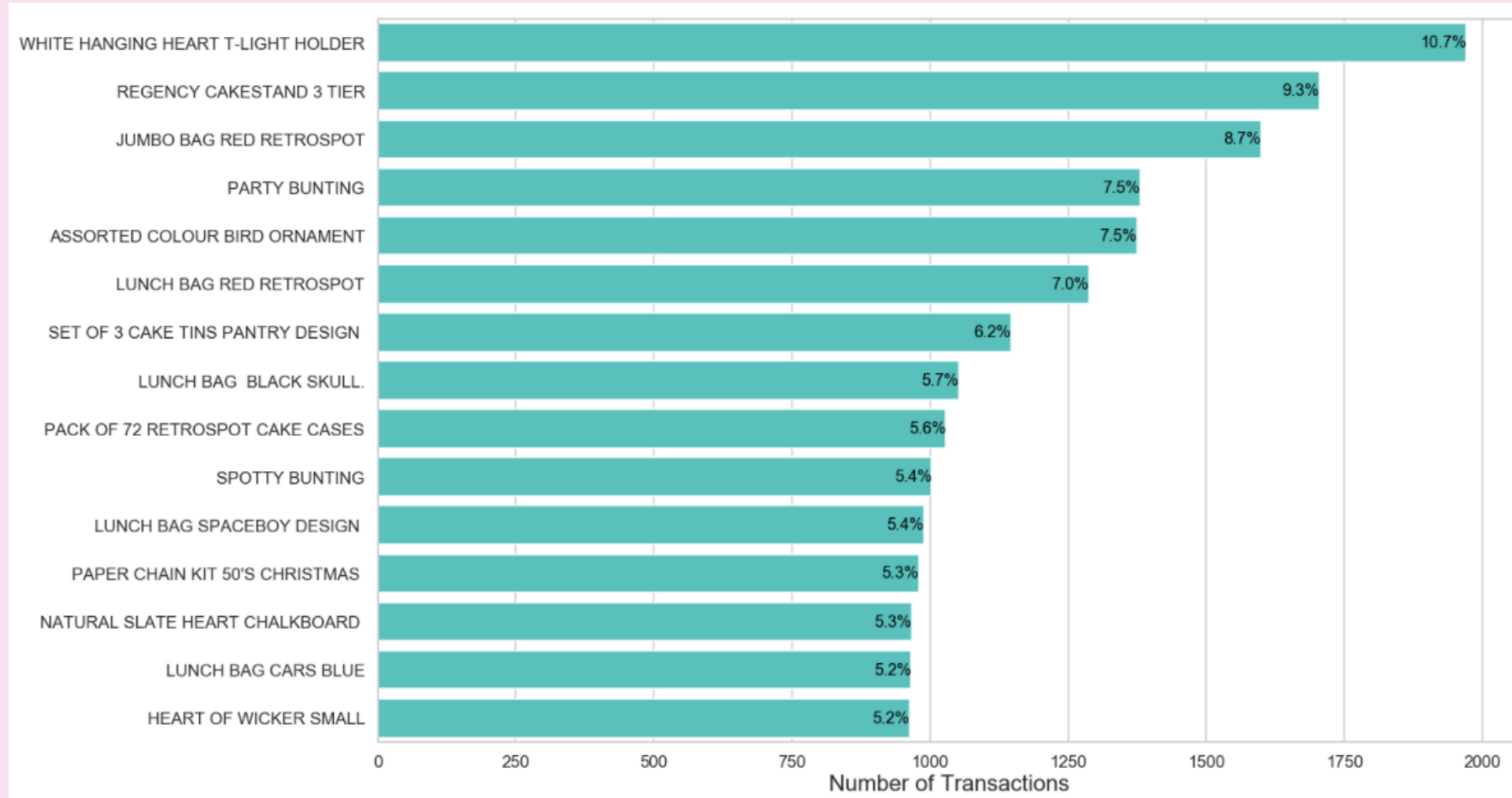
# Long Tail Plot



- Long Tailed Distribution
- Most products are in the “long tail”
- Only a small percentage of products have a high volume of transactions
- The 2002 tail products only represent 10% of the transactions

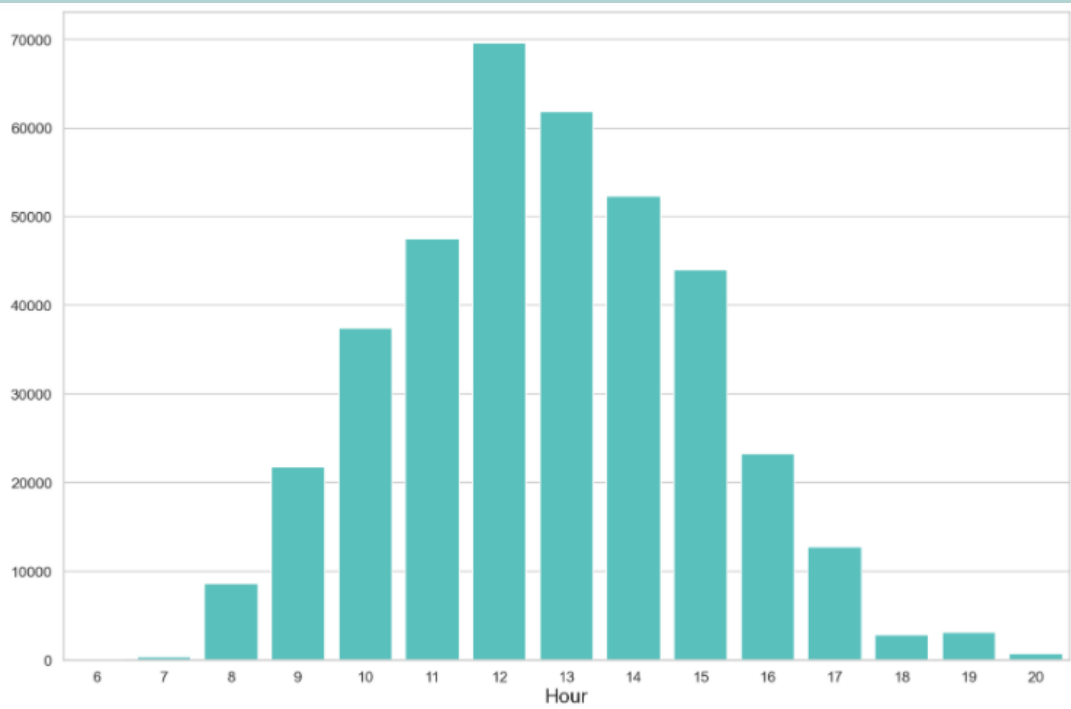
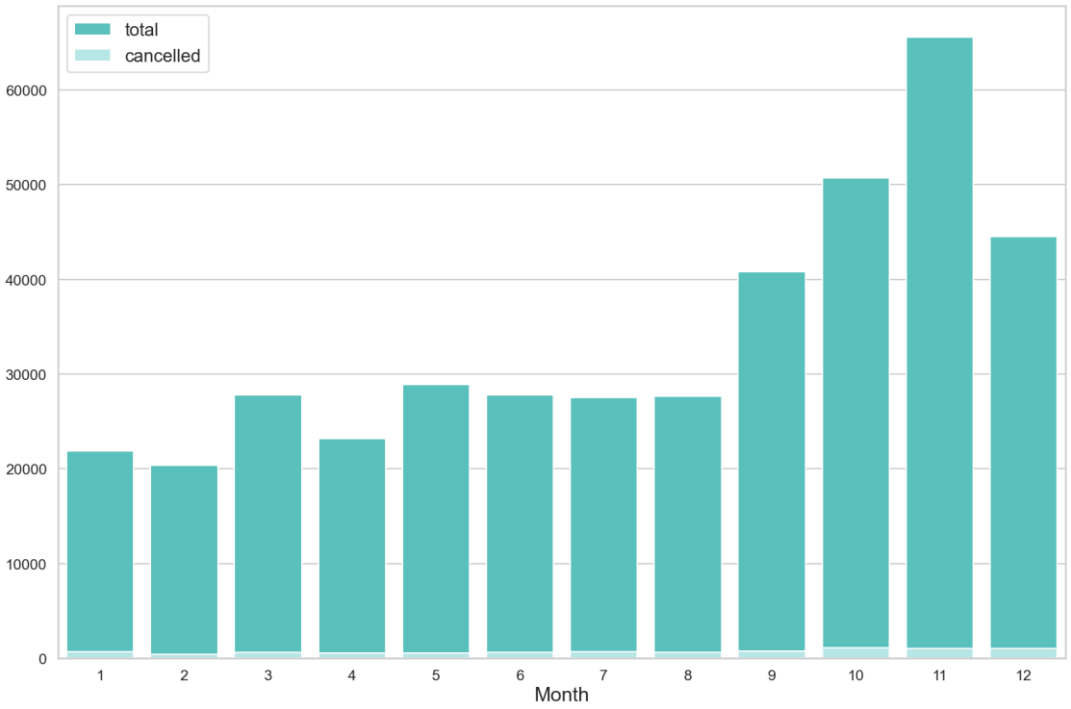
## INSIGHTS

## ZOOM: TOP 15 products



# INSIGHTS

# Number of Transactions Over Time

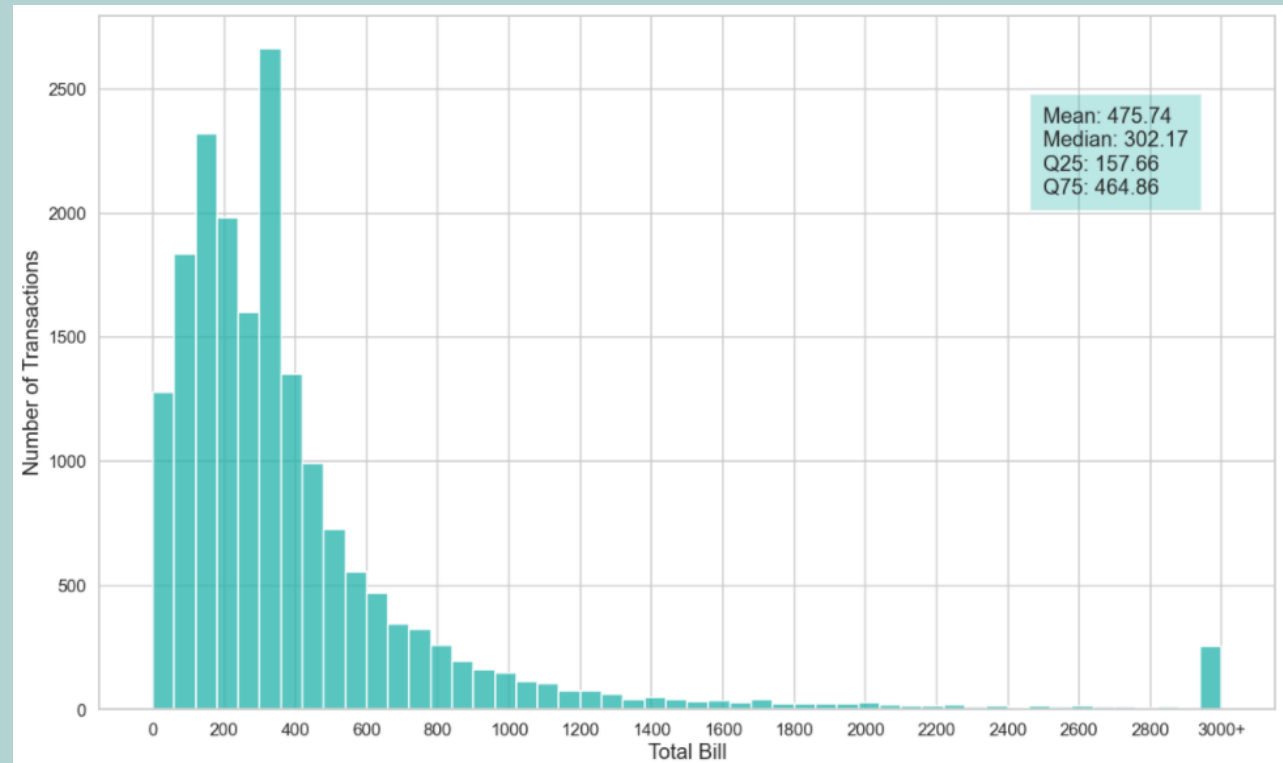


INSIGHTS



## Transaction Bill

- Skewed distribution
- Most frequent spend amount: 300\$ to 350\$
- 75% of the transactions have a bill lower than \$464.86
- Max: 168 469.6\$



# INSIGHTS

# RECOMMENDER SYSTEMS

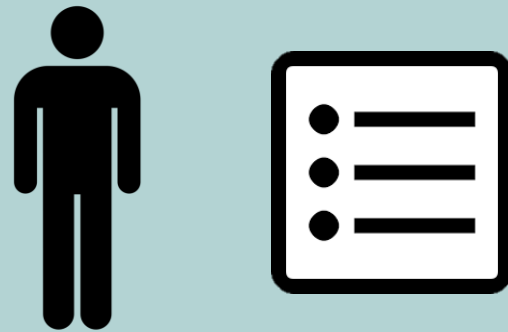
## Additional Products of Interest

ALS Model



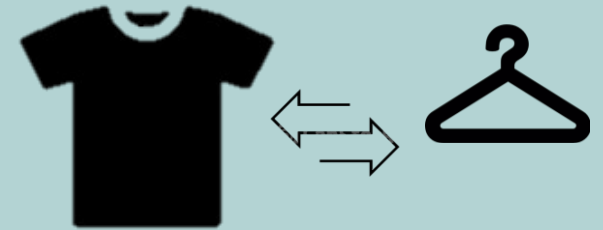
## Your First Time

Popular Recommender Model



## Fill Your Basket

Apriori Algorithm



# RESULTS

## ALS Model

- K = 5 Products

	pop_model	als_model
precision	0.095360	0.060253
map	0.061715	0.036538
ndcg	0.100412	0.060119
auc	0.506615	<b>0.504074</b>
personalization	NO	YES

## Popular Model

- K = 10 Products

	pop_model
precision	1.000000
map	1.000000
ndcg	1.000000
auc	<b>0.501992</b>
personalization	NO

## Association Rules

- Lift > 1
- $0.510 < \text{Confidence} < 0.757$

support	confidence	lift
0.041301	0.756650	13.198791
0.041301	0.720450	13.198791
0.044367	0.677340	6.020613
0.038935	0.611486	5.435273
0.036569	0.578723	5.144054
0.035225	0.509728	6.056476

# DEPLOYMENT



Personalized menu,  
products' categories



Products ratings  
and reviews



Layout of landing  
page with  
preferences  
at the top



Coupons, bundle  
promotions

**Thank You!**

