

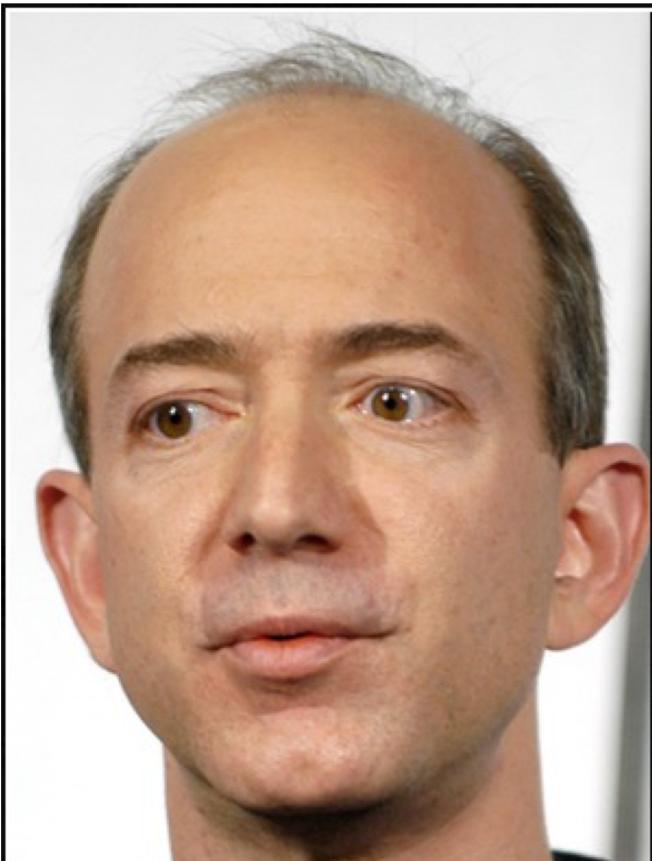
Customer Segmentation

for an e-commerce retailer from the UK



Start with WHY

WHY Customer Segmentation?



We see our customers as invited guests to a party, and we are the hosts. It's our job every day to make every important aspect of the customer experience a little bit better.

— Jeff Bezos —

AZ QUOTES

Background & Problem Statement

About the dataset:

Source: Kaggle

Date: Between Dec 2010 and Dec 2011

UK-based and non-store online retail.

Mainly sells unique all-occasion gifts.

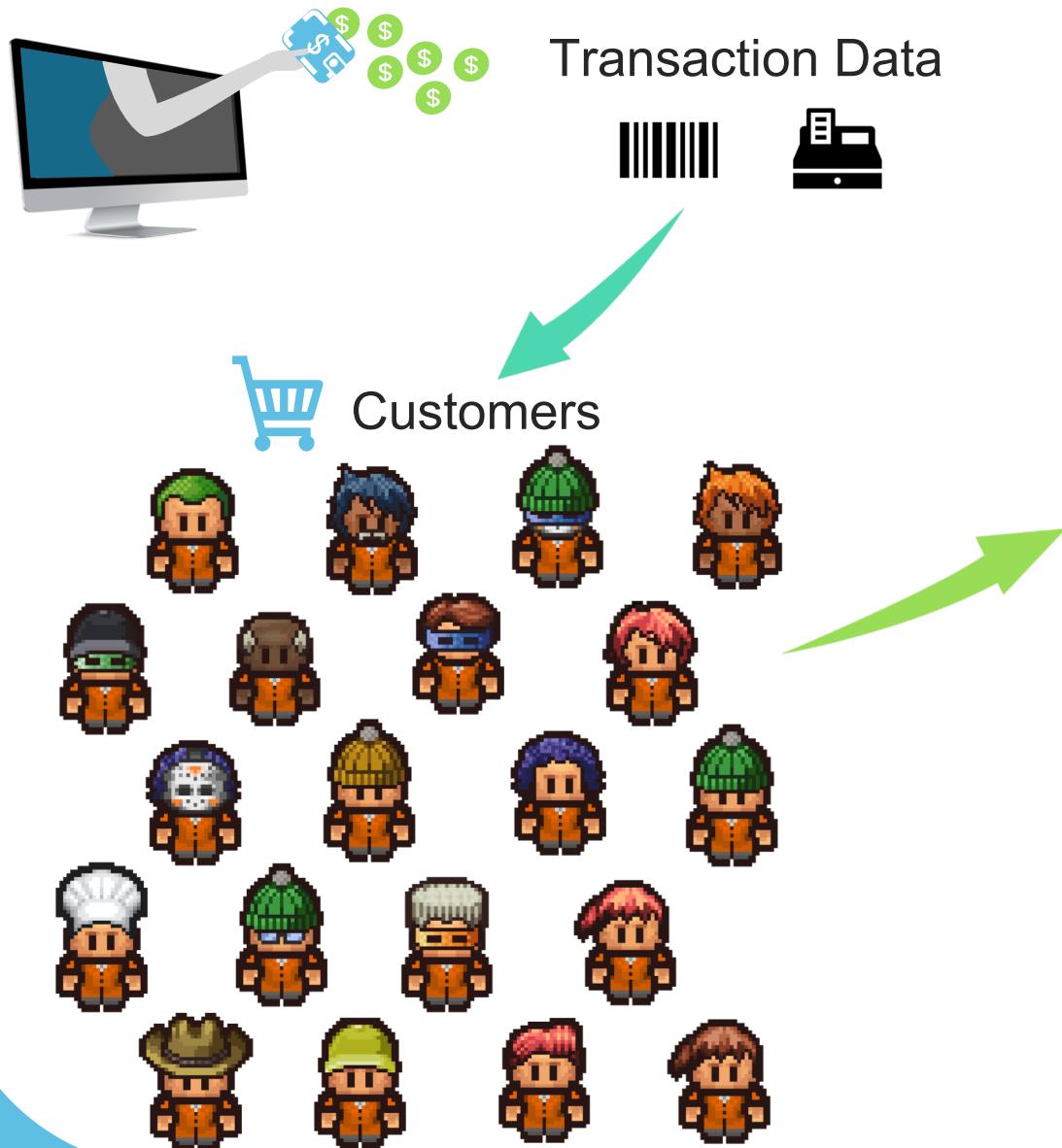
Most customers of the company are wholesalers.

Problem Statement:

Group and understand the habits of the customers so that the company can gain a deeper understanding of their preferences.



OBJECTIVE:



Loyal customers



Heavy occasional customers



Light loyal customer



Occasional customers



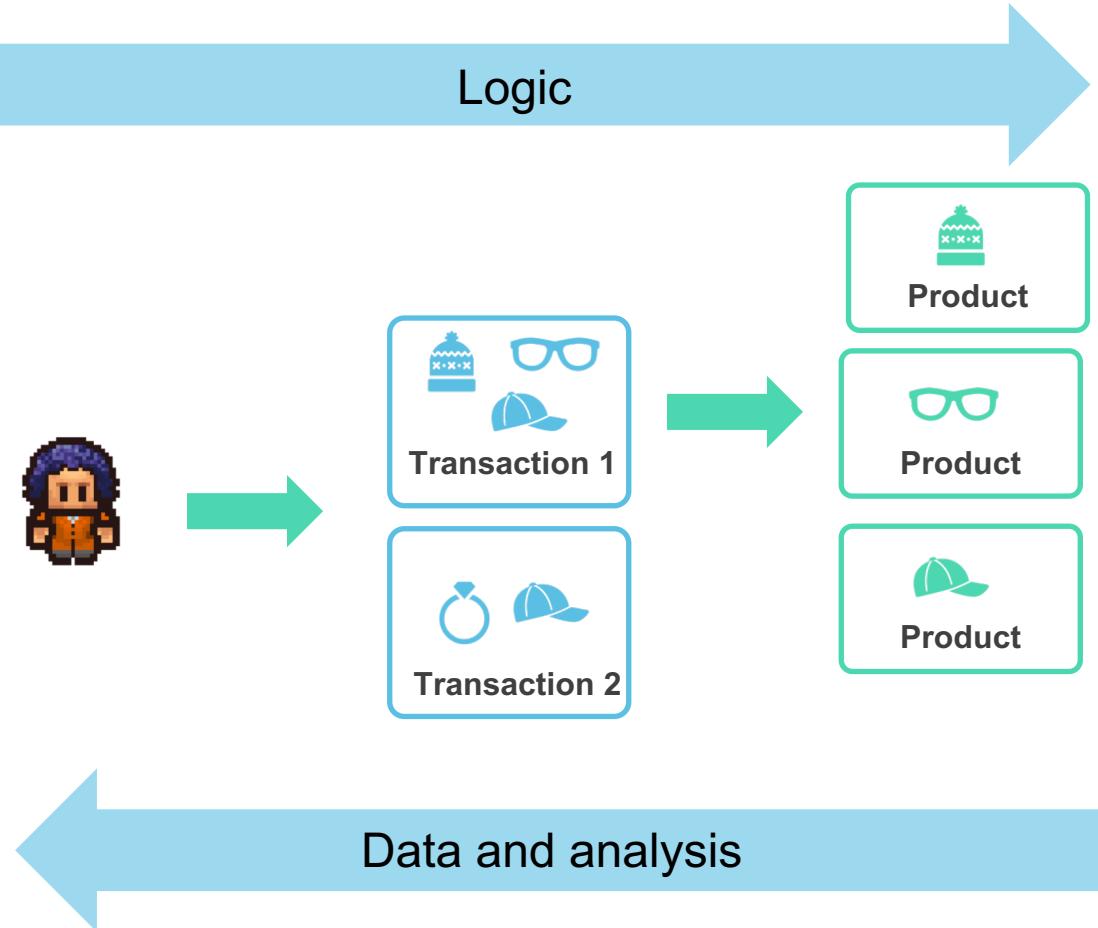
Trialist



01

Data Exploration and Findings

Concepts and Terminologies



01

Data Exploration and Findings

Some snapshots of the data



Items purchased:

541,909



Most popular quantity of purchase and median spend on each item:
1 - \$10-\$15



Max Quantity in one transaction:

4,800



Unique Customer ID:

5,703



Cancelled Transactions:

9,725



Median Spend per transaction

\$300



Highest transactional Month

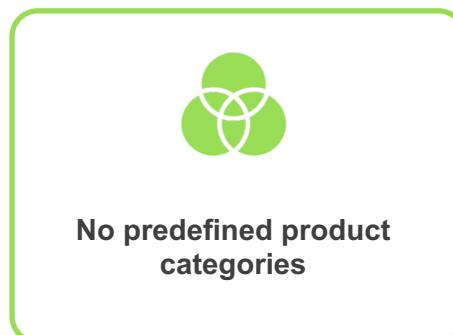
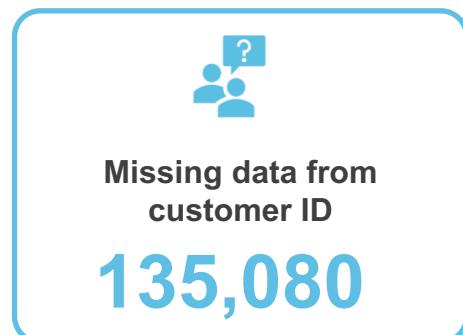
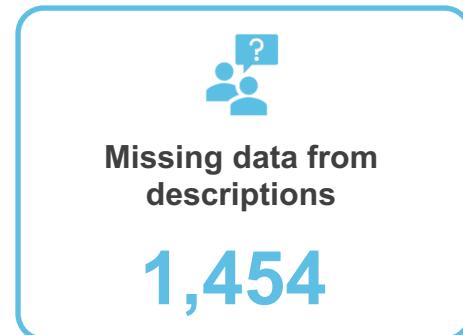
November

Top 10 Transactions by country

Country	Transaction
United Kingdom	488843
Germany	9480
France	8541
EIRE(aka Ireland)	8184
Spain	2528
Netherlands	2371
Belgium	2069
Switzerland	1994
Portugal	1510
Australia	1258

02

Assumptions, Risks and Limitations



InvoiceNo	StockCode	Description	Qu
540673	21644	found	
547673	47593A	found	
550950	17091A	found	
556963	35965	found some more on shelf	
558369	21082	found	
558725	84944	found	
559335	21147	found	
560329	35965	found	
567129	72127	found box	

03

Clustering of products into categories



Description



Description



Description



Description



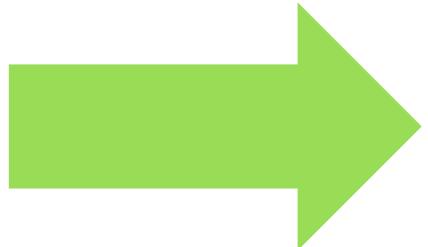
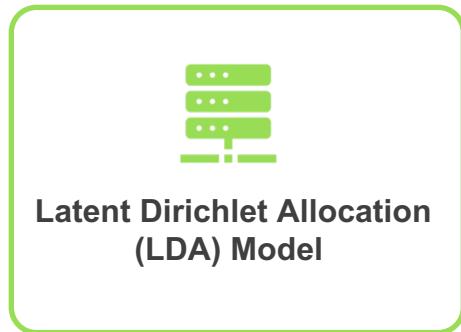
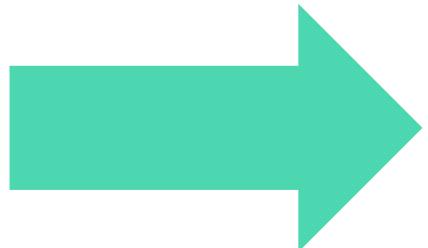
Description



Description



Description



Category A



Category B



Category C

03

Clustering of products into categories

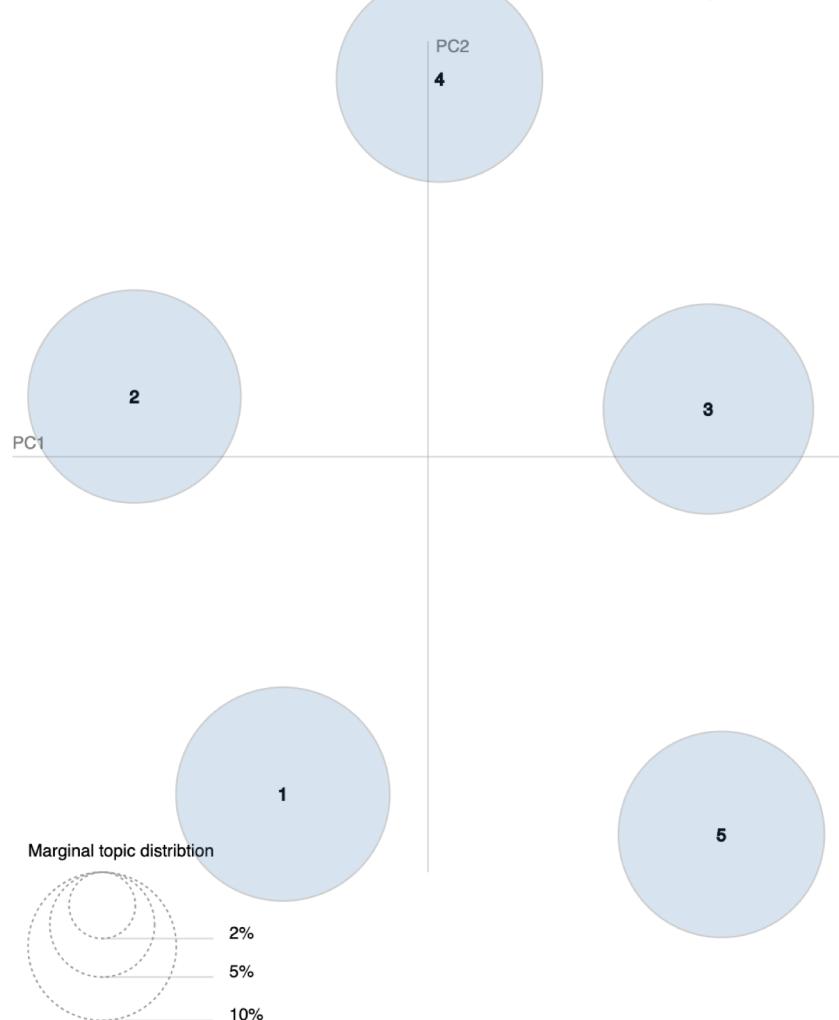
Selected Topic: 0 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:(2)

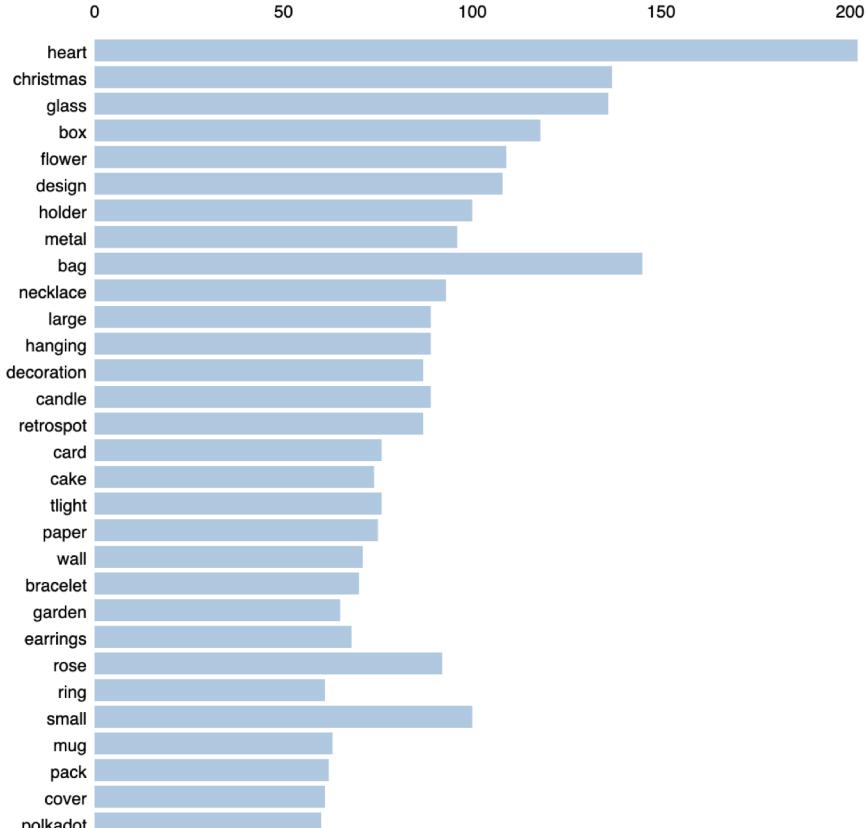
$\lambda = 1$



Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Salient Terms¹



Overall term frequency

Estimated term frequency within the selected topic

1. $\text{salience}(\text{term } w) = \text{frequency}(w) * [\sum_t p(t|w) * \log(p(t|w)/p(t))]$ for topics t; see Chuang et. al (2012)

2. $\text{relevance}(\text{term } w | \text{topic } t) = \lambda * p(w|t) + (1 - \lambda) * p(w|t)/p(w)$; see Sievert & Shirley (2014)

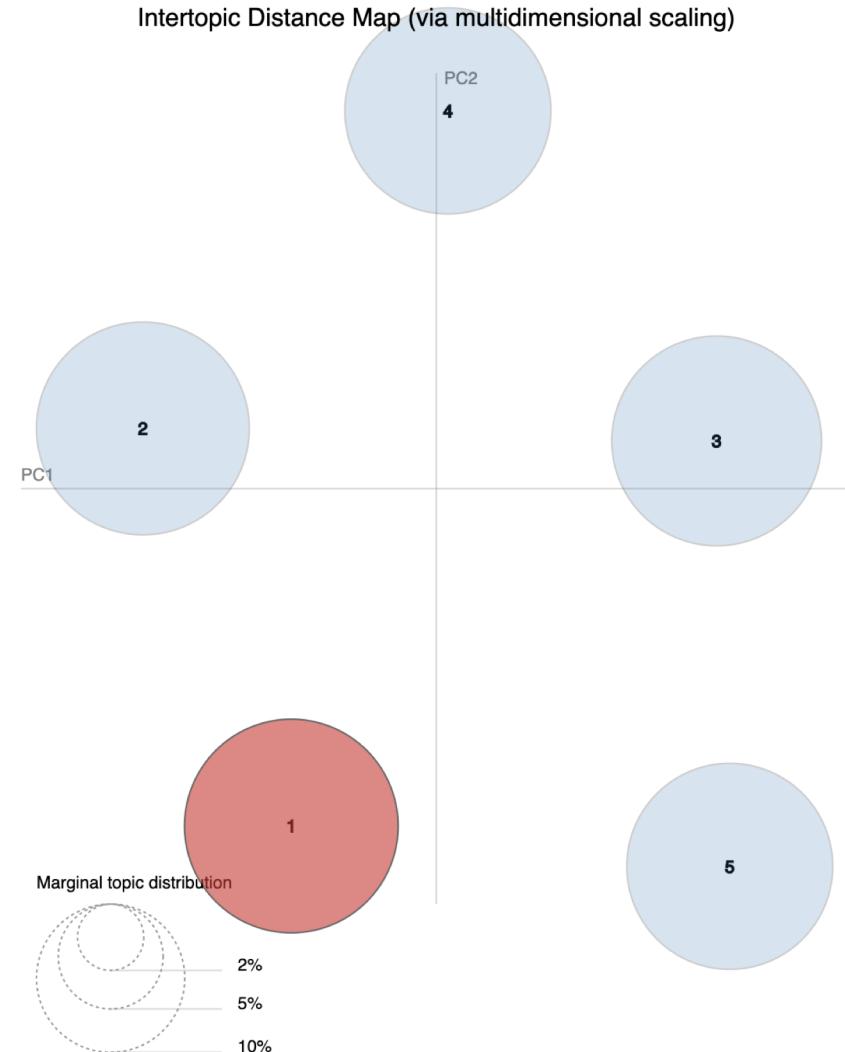
03

Clustering of products into categories

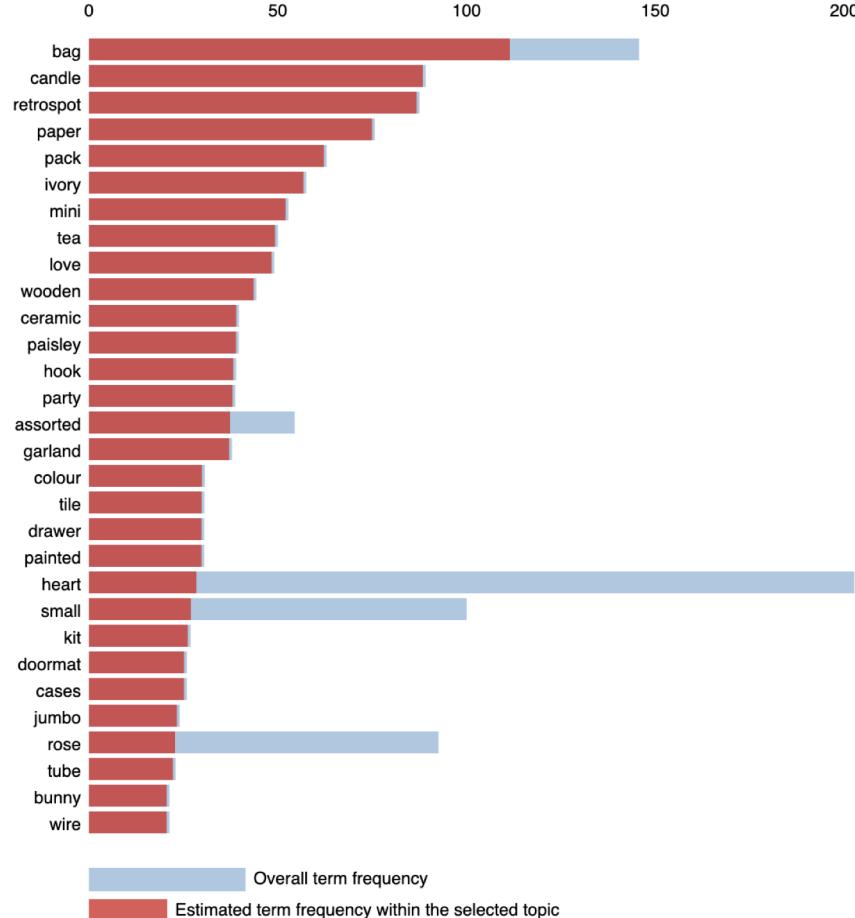
Selected Topic: 1 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:(2)

$\lambda = 1$



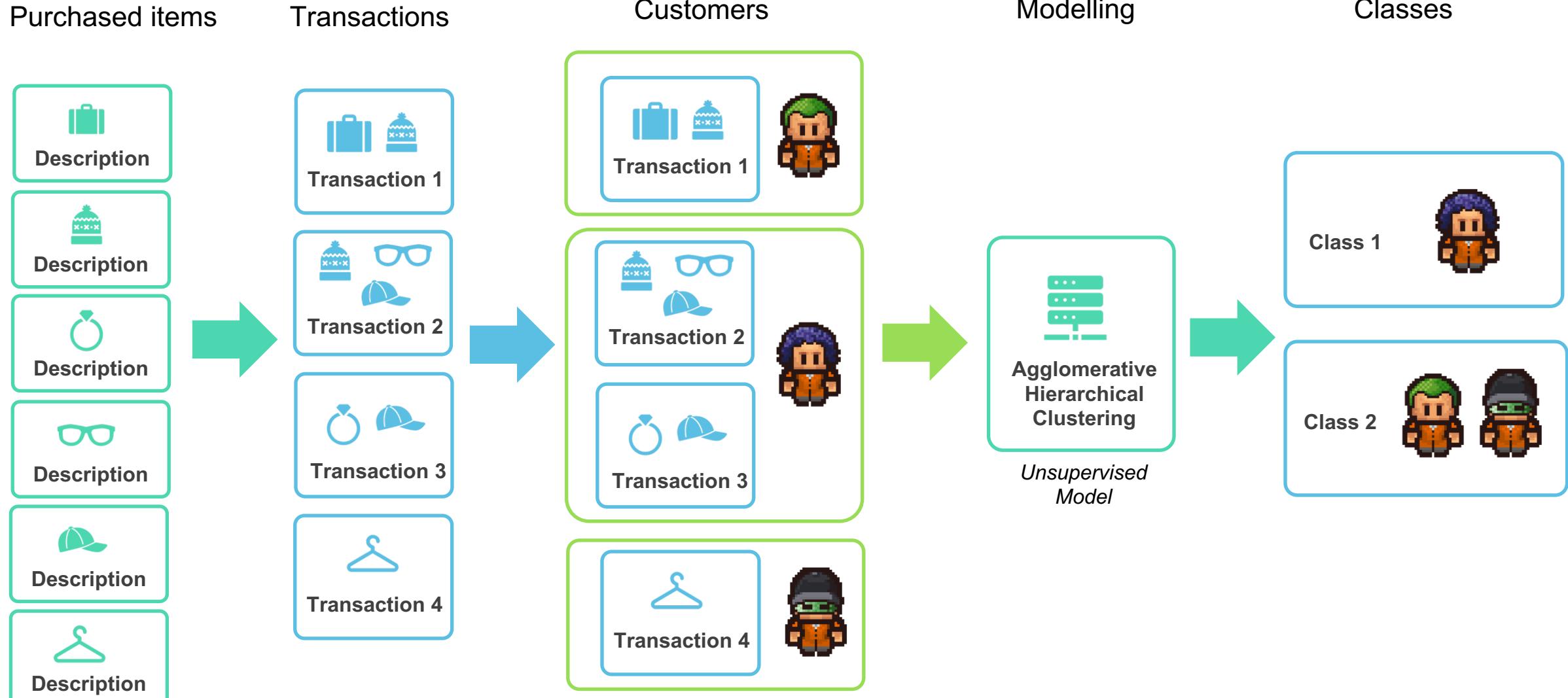
Top-30 Most Relevant Terms for Topic 1 (20.8% of tokens)



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

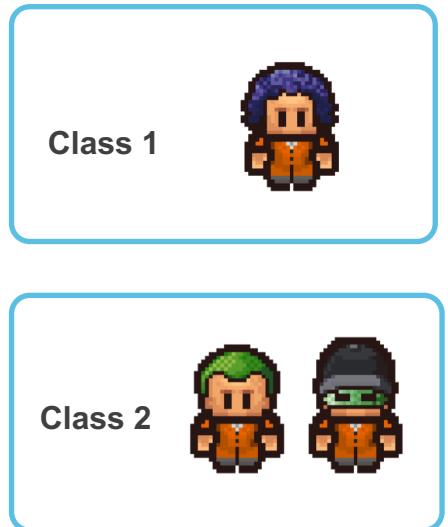
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t) / p(w)$; see Sievert & Shirley (2014)

Grouping customers into classes

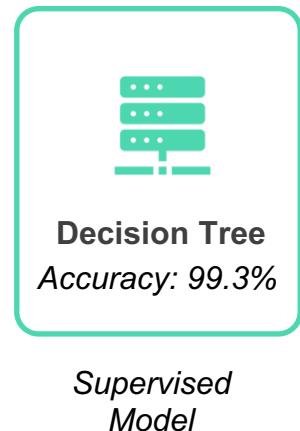


Grouping customers into classes

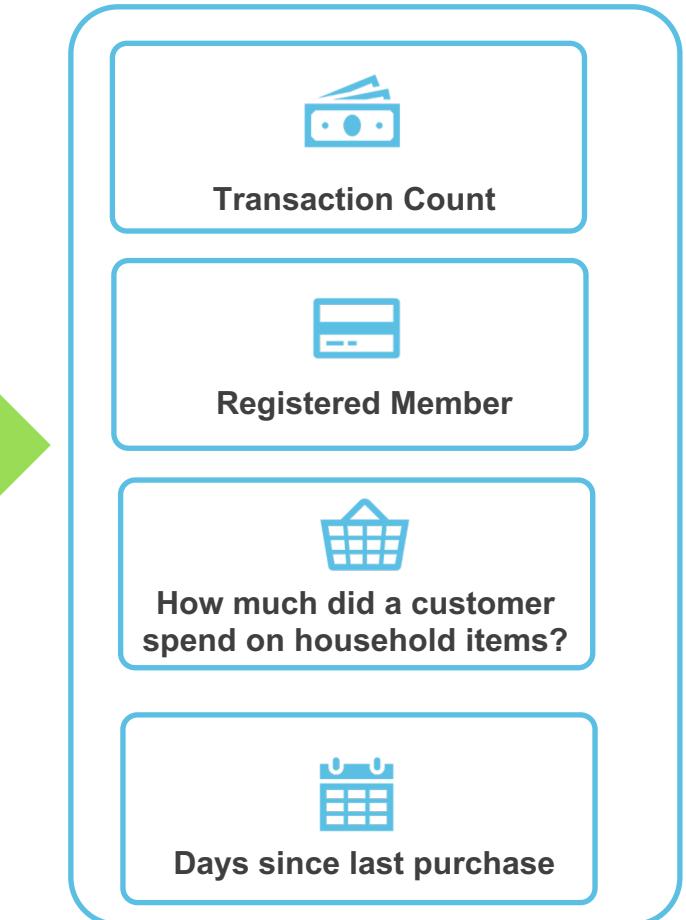
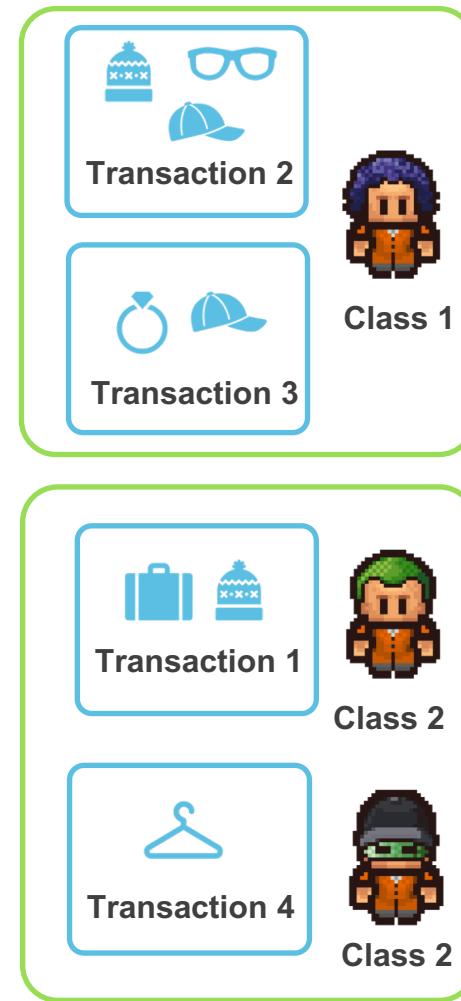
Classes



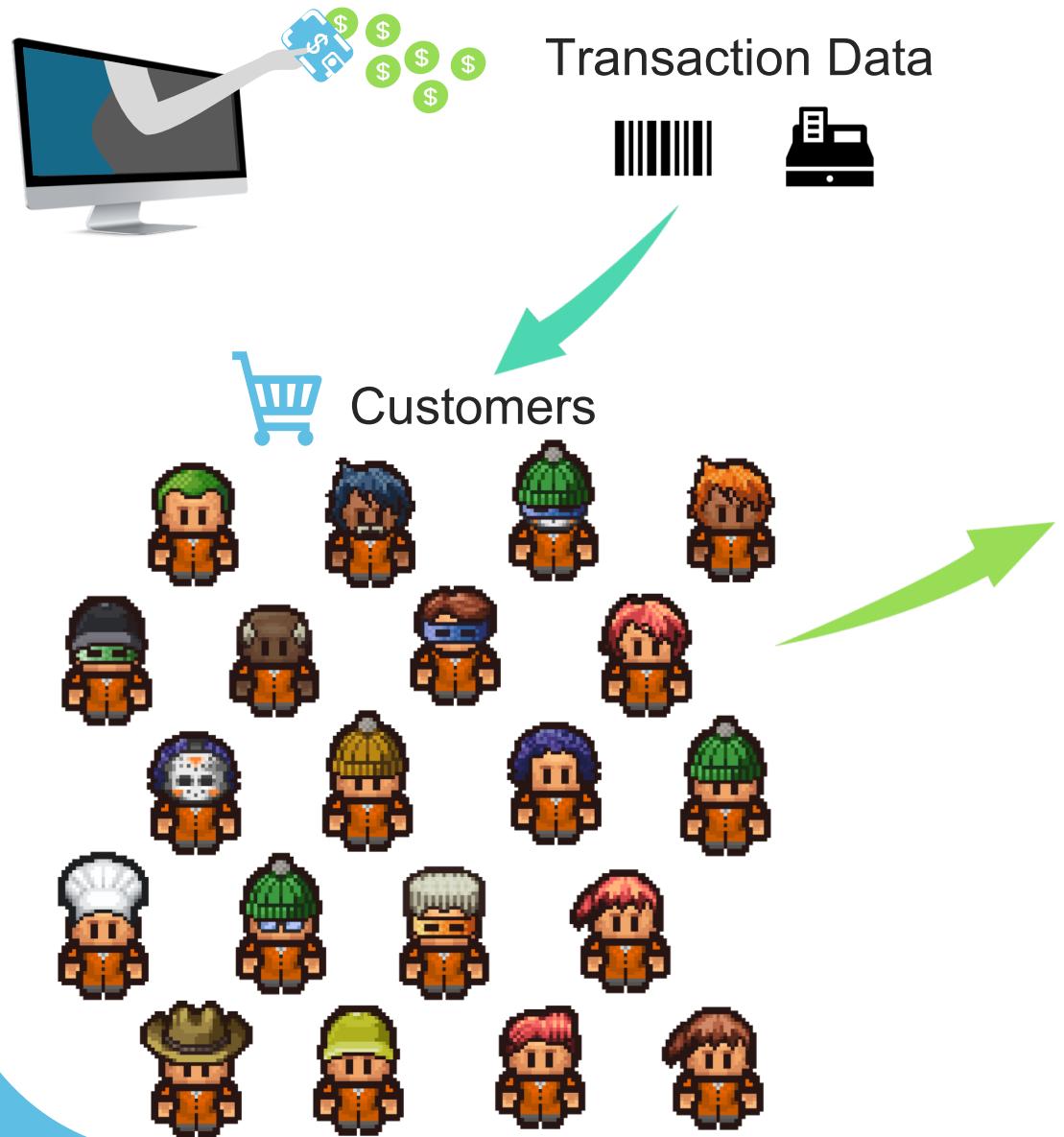
Modelling



Characteristics of each class



Review:



Loyal customers



Heavy occasional customers



Light loyal customer



Occasional customers



Trialist



05

Understand their habits

Loyal customers

Heavy occasional customers

Light loyal customer

Occasional customers

Trialist

High frequency, high spend



Number of Customers

8



Transaction Count

Minimum 17



Registered Members?

100%



Average Quantity Per Transaction (Median)

1,250



Average Spend Per Transaction (Median)

\$3,500



Total Spendings (Median)

\$128,000



Average Days Per Invoice (Median)

6 Days



Days Since Last Purchase (Median)

2.5 Days

05

Understand their habits

Loyal customers

Heavy occasional customers

Light loyal customer

Occasional customers

Trialist

Number of Customers

1187



One time purchase

Transaction Count

1



Registered Members?

0%



Average Quantity Per Transaction (Median)

57



Average Spend Per Transaction (Median)

\$247



Total Spendings (Median)

\$247



Average Days Per Invoice (Median)

0 Days



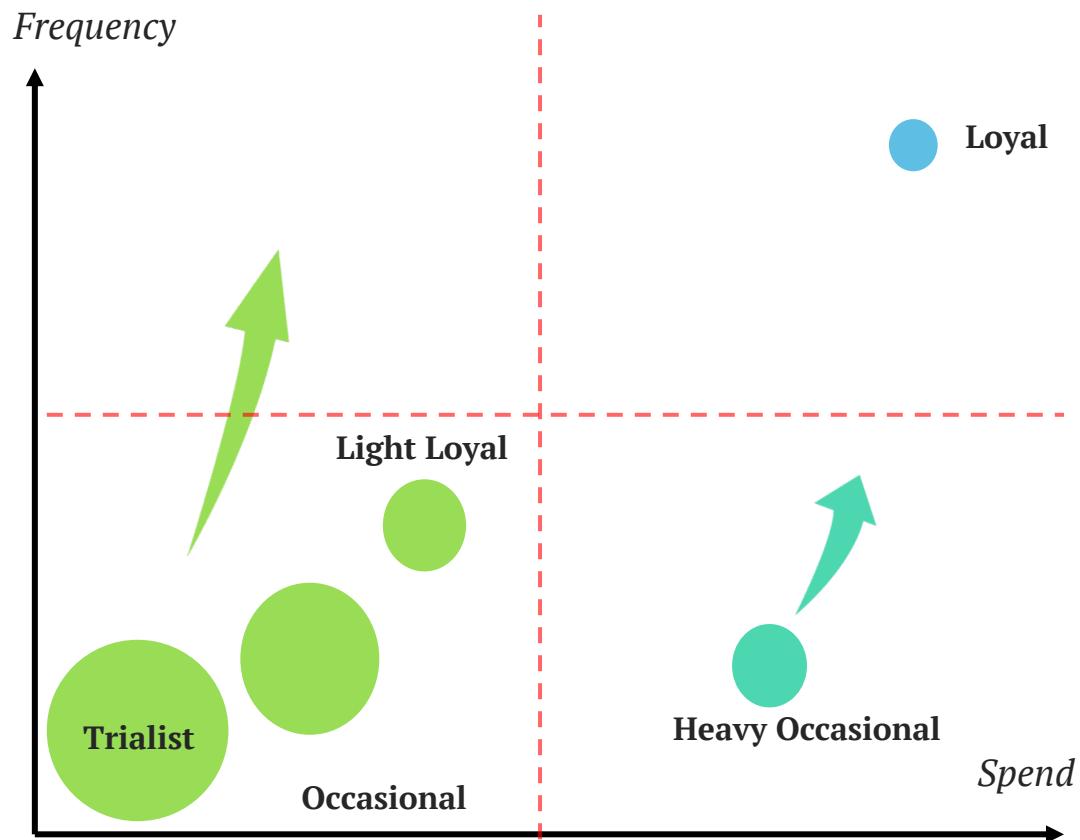
Days Since Last Purchase (Median)

200 Days

Moving Forward

Spend – Frequency Matrix

• The size indicates the group size



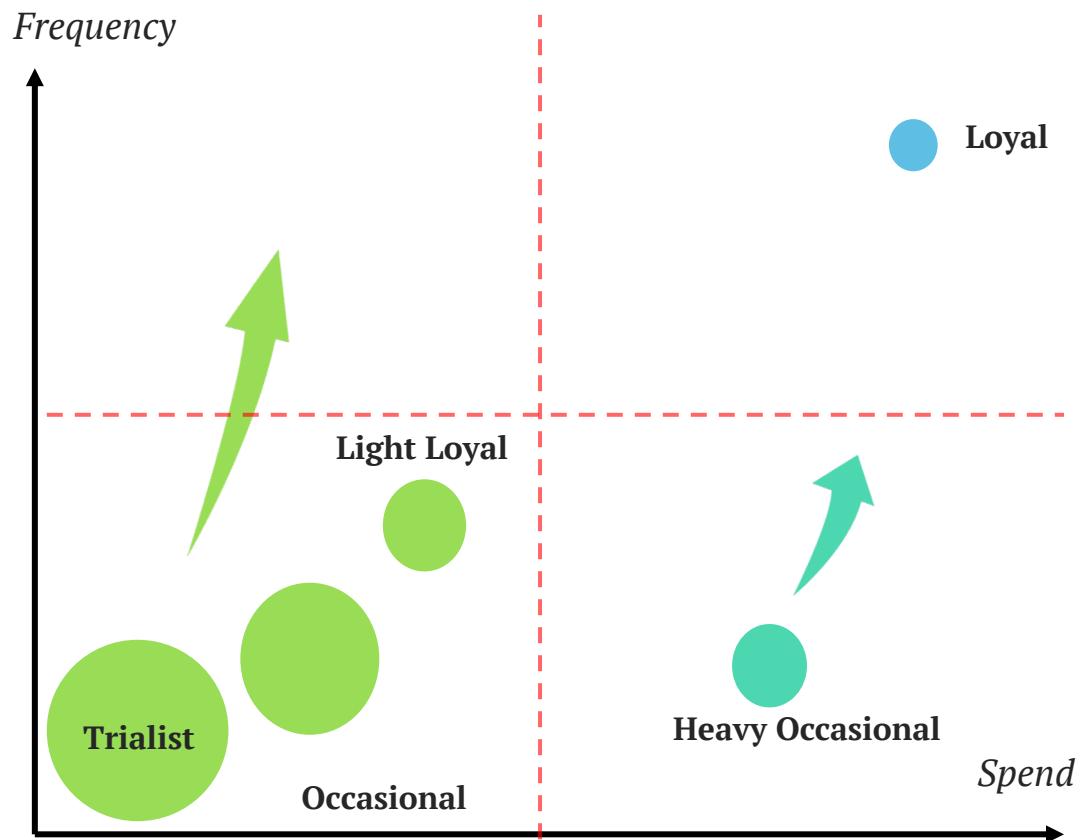
With the customers segmented, we are able to better:

- Personalise our messages to our customers
- Make more effective acquisition and retention
- Better ROI from Marketing
- Unearth New Opportunities

Recommendations

Spend – Frequency Matrix

• The size indicates the group size



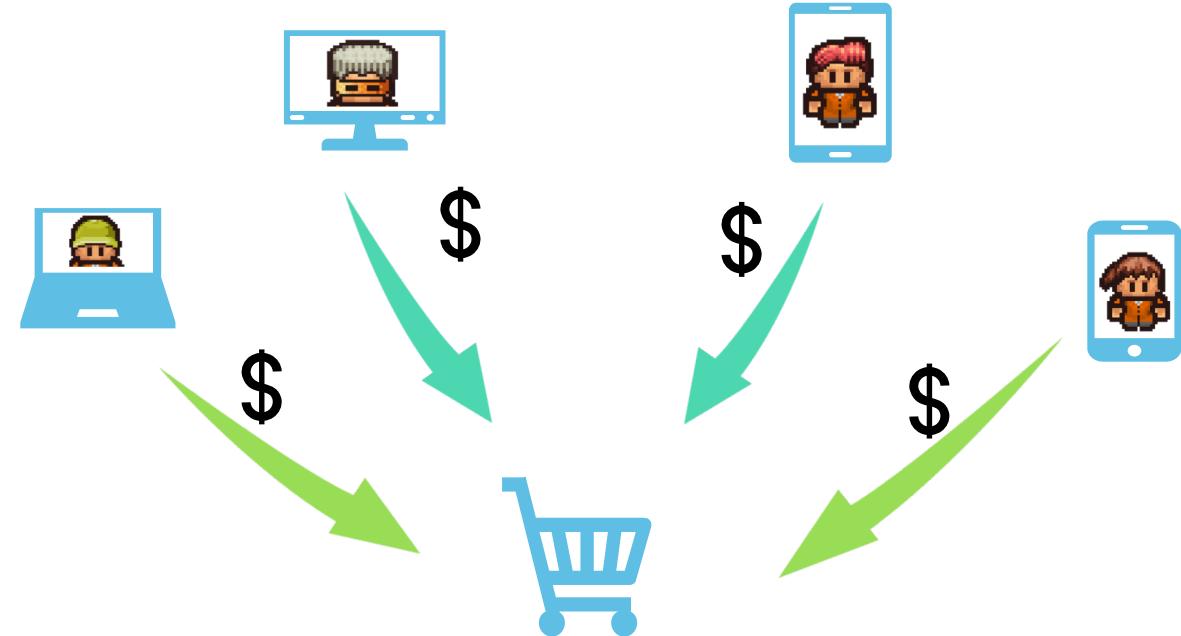
- Acquire trialist – offer a discount code or free delivery upon signing up
- Occasional / Light loyalist - offer a return discount code for their next purchase to increase frequency
- Heavy occasional – recommend products that are relevant to them so that they can buy more from the site
- Loyalist – Create a rewards program so that we can retain them and not lost them to our competitors

So What??

How does these actions benefit the business??

	Number of Customers	Average Spend Per Transaction (Median)	Potential Incremental
Loyal customers	8 4	x \$3,500	= \$14,000
Heavy occasional customers	192 96	x \$3,200	= \$307,200
Light loyal customer	447 223	x \$380	= \$84,740
Occasional customers	3869 1934	x \$275	= \$531,850
Trialist	1187 593	x \$247	= \$146,471
			\$1,084,261

11% of
annual
revenue!



Customer Segmentation

for an e-commerce retailer from the UK

Any Questions?

<https://cw-customer-segmentation.herokuapp.com/>