

Machine Learning for

# CUSTOMER SEGMENTATION

Descriptiv & Predictive  
Customer Analytics

We don't know our customers and how to target them!

**NEED:**

Marketing & Sales Management of e-commerce retailer "Deco Word"

need a high quality segmentation of customers  
on their shopping behaviour

in order to tailor specific targeting

# Data Set Overview



Online retail transactions between 01/12/2010 and 09/12/2011 for a UK-based online shop. The company mainly sells unique all-occasion gifts.

## CUSTOMER:

1. Customer ID
2. Country

## ORDERS:

1. Invoice Number
2. Invoice Date
3. Quantity

## PRODUCT:

1. Stock Code
2. Description
3. Unit Price

- 541.909 Instances
- 8 features
- 20% missing values in customer ID
- Aggregation of transactions by customer
- instances ready for clustering: 3.526

# Business Insights



Total Revenue  
8.370.632 £



Customer  
4.327

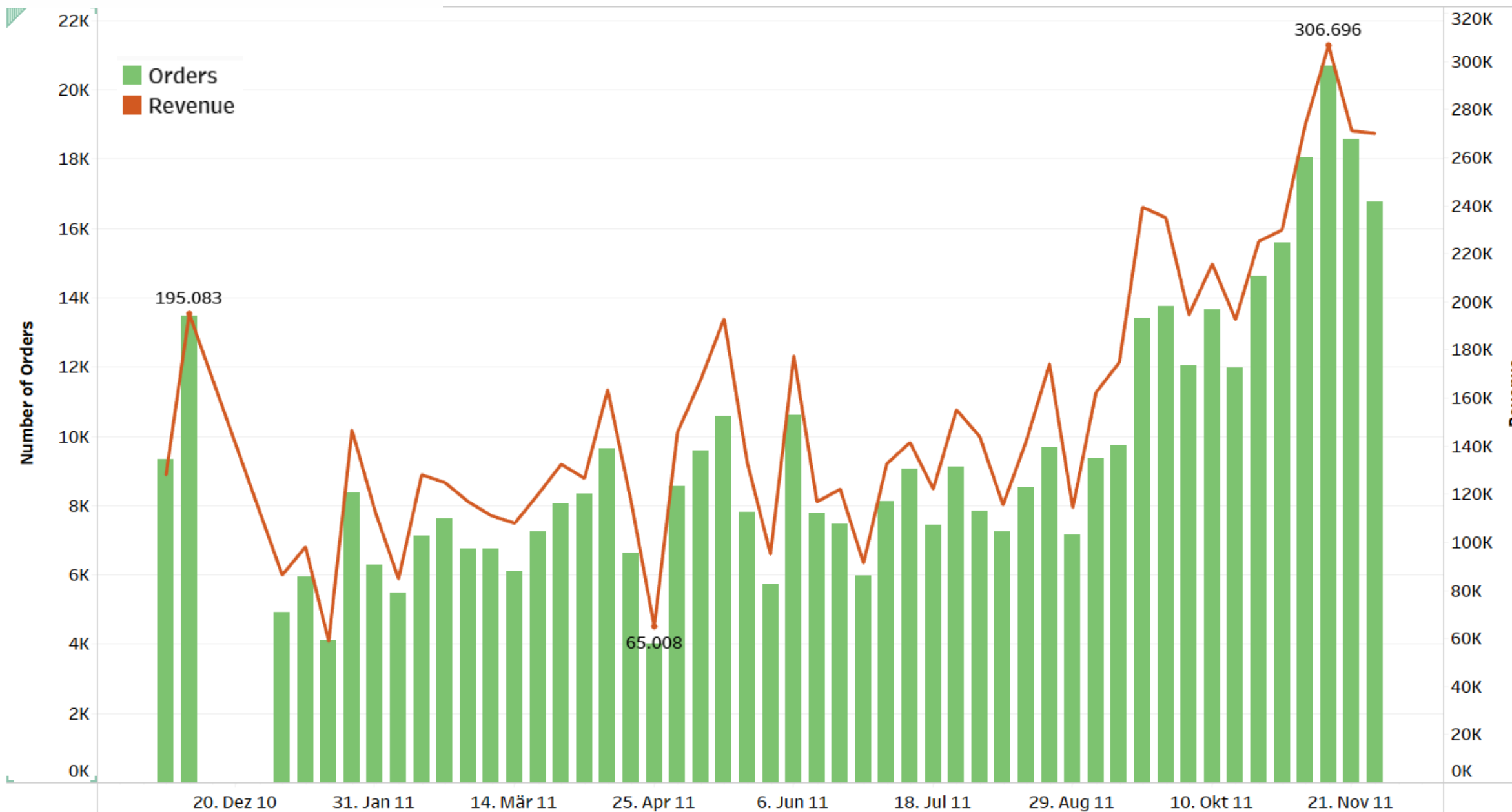


Transactions  
18.456

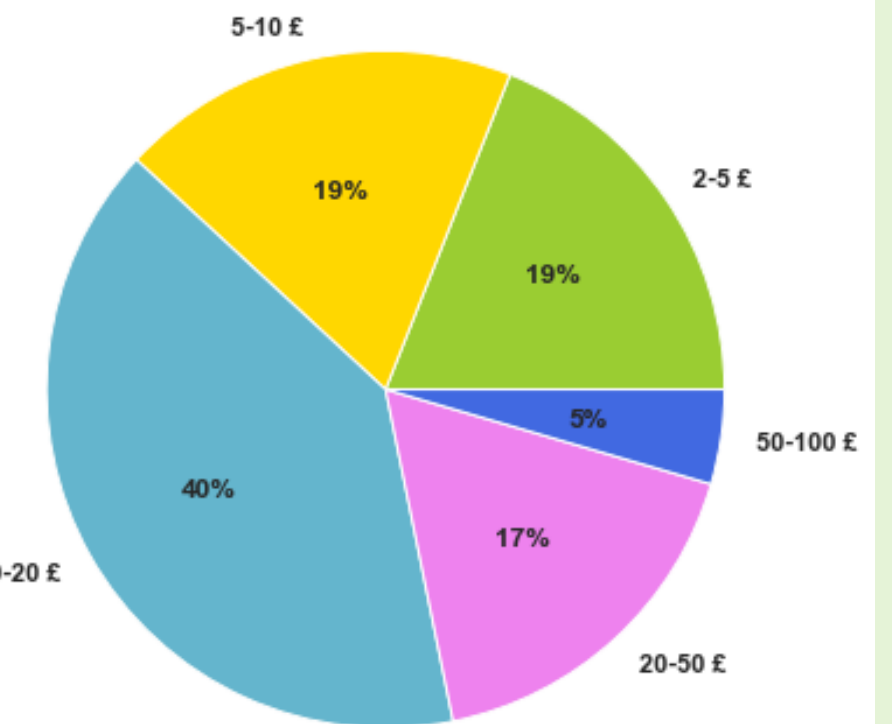


Products  
3684

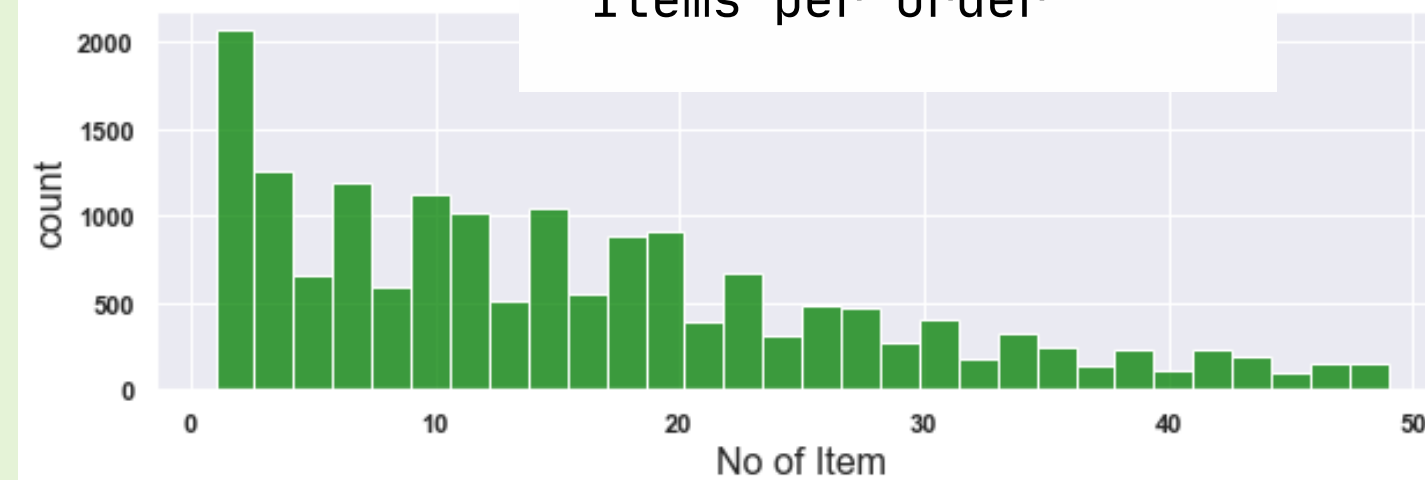
## Revenue & Orders



## Value per Order



## Items per Order



# Market Basket Analysis

with the association rule

	antecedents	consequents	confidence	lift
0	(ALARM CLOCK BAKELIKE PINK)	(ALARM CLOCK BAKELIKE GREEN)	0.725000	7.421711
18	(SET/6 RED SPOTTY PAPER CUPS)	(SET/6 RED SPOTTY PAPER PLATES)	0.888889	6.915556

Customer 7 times more likely buys green alarm, when bought the pink  
\*in 72% of all cases

Customer 6 times more likely buys paper plates, when bought paper cups  
\*in 88% of all cases

## INSIGHT:

- recommend products in different variations & colors or sets
- nurture basket size by bulk discounts

# RFM Segmentation

R

RECENCY - of last order

M

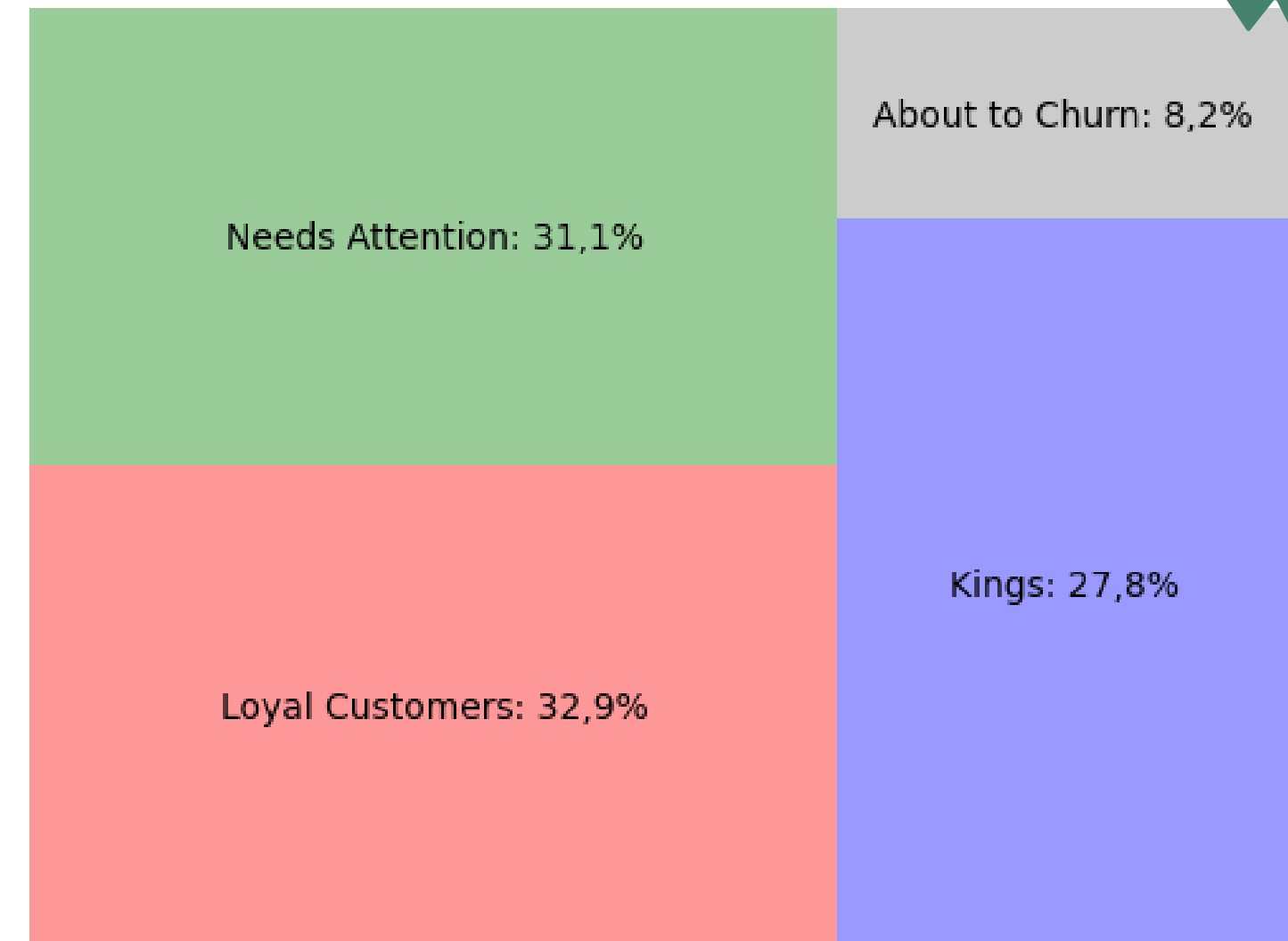
MONETARY - total purchase amount

F

FREQUENCY - number of orders

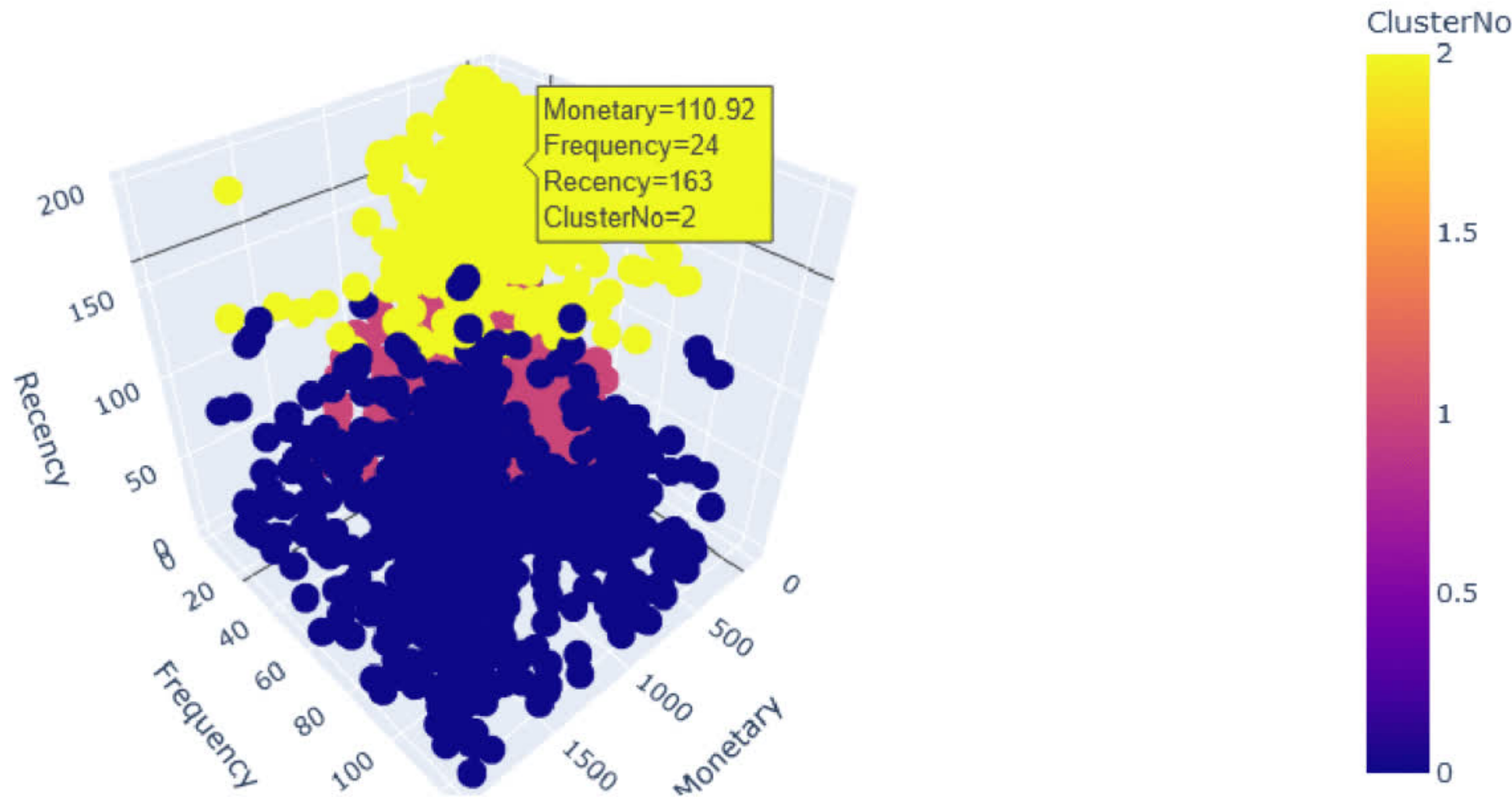
## RFM Score - segmentation

### RFM Segments



# K Means Clustering based on RFM features

## UK Customer Cluster



	Monetary	Frequency	Recency
ClusterNo			
0	1181.29	73.72	41.47
1	405.71	24.82	41.25
2	415.38	25.85	150.57

# Insights UK Customer Cluster

## cluster 0

31% of all customer

**M**

- 1.181 £ total spend (cluster mean)

**F**

- 73 orders per year  
= 1 order per week/average

**R**

- last order 1,5 months ago

> most valuable cluster  
loyal & hot customer

## cluster 1

49% of all customer

- 406 £ total spend
- 17 £ average order value

- 25 orders per year  
= order every 2 weeks/average

- last order 1,5 months ago

> loyal & hot customer,  
but less value

## cluster 2

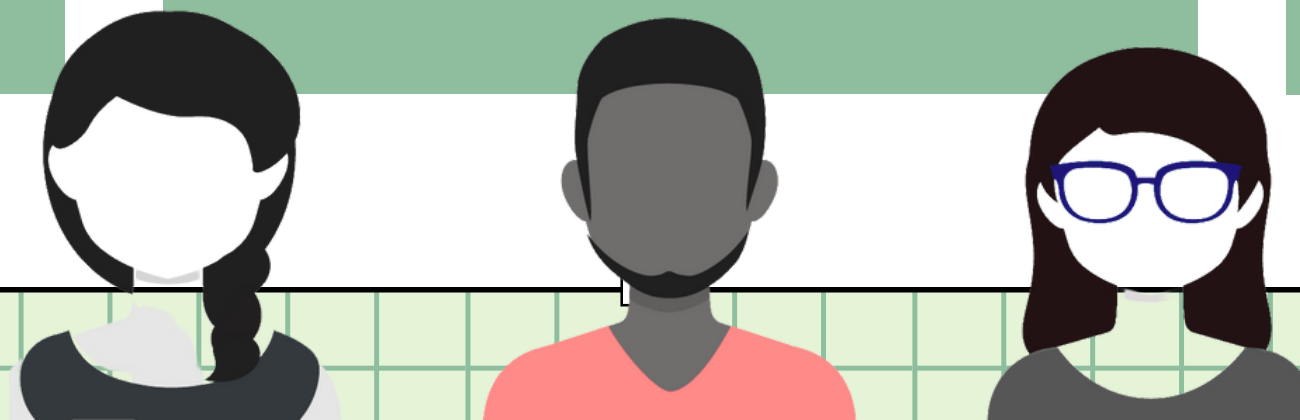
20% of all customer

- 415 £ total spend

- 26 orders per year  
= order every 2 weeks/average

- last order 5 months ago

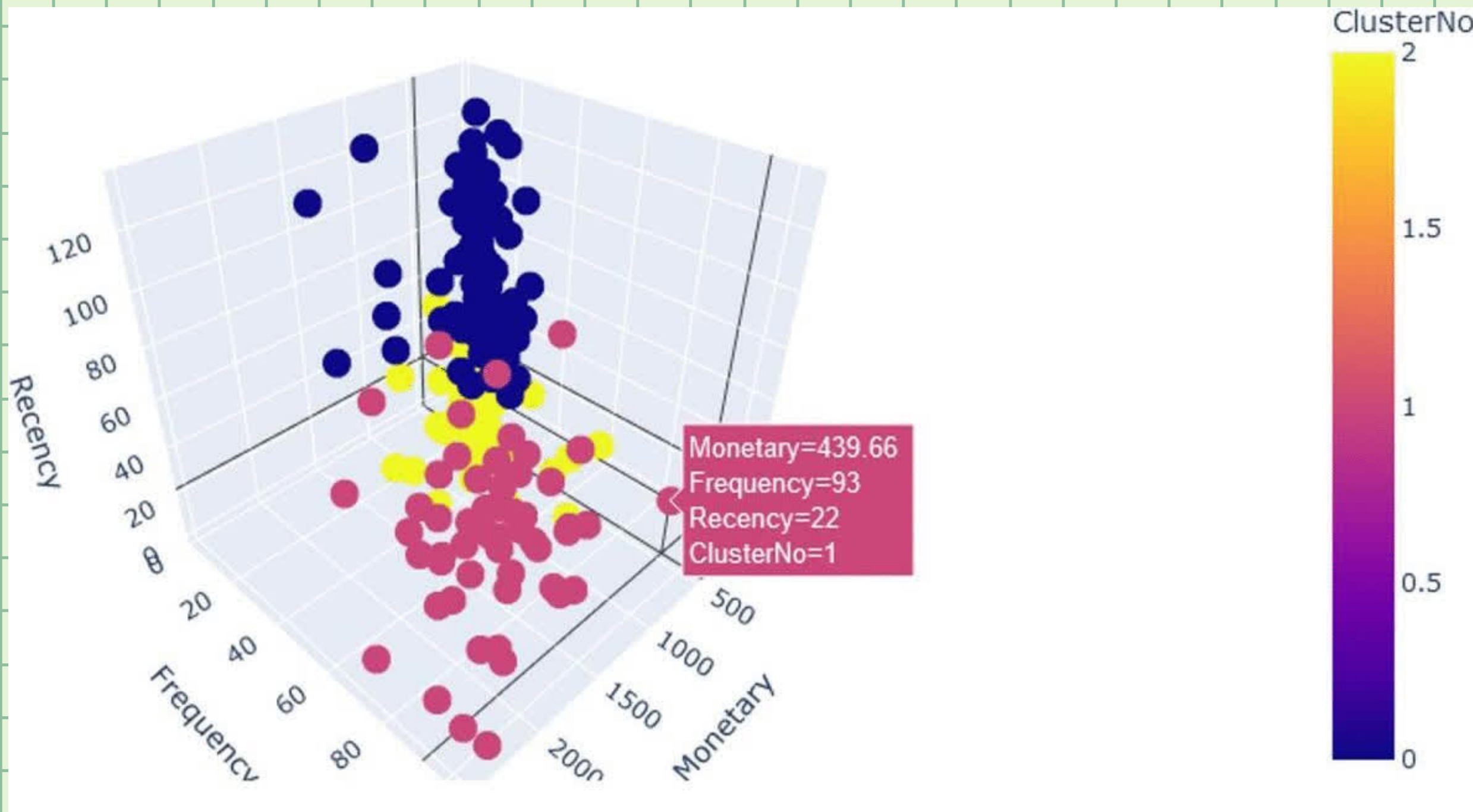
> needs attention,  
about to churn!





# World Customer Cluster

K Means Clustering based on RFM features



	Monetary	Frequency	Recency
ClusterNo			
0	689.83	31.27	86.83
1	1599.43	74.81	34.85
2	541.55	27.83	21.96

# Insights World Customer Cluster

## cluster 0

36,5% of all customer

M

- 690 £ total spend  
(cluster mean)

F

- 31 orders per year  
= order every 3 weeks/average

R

- last order 3 months ago
- > valuable, less active

## cluster 1

27% of all customer

- 1.600 £ total spend
- 22 £ average order value

- 75 orders per year  
= order every 5 days/average

- last order < 1 months ago

> most valuable cluster  
loyal & hot customer

## cluster 2

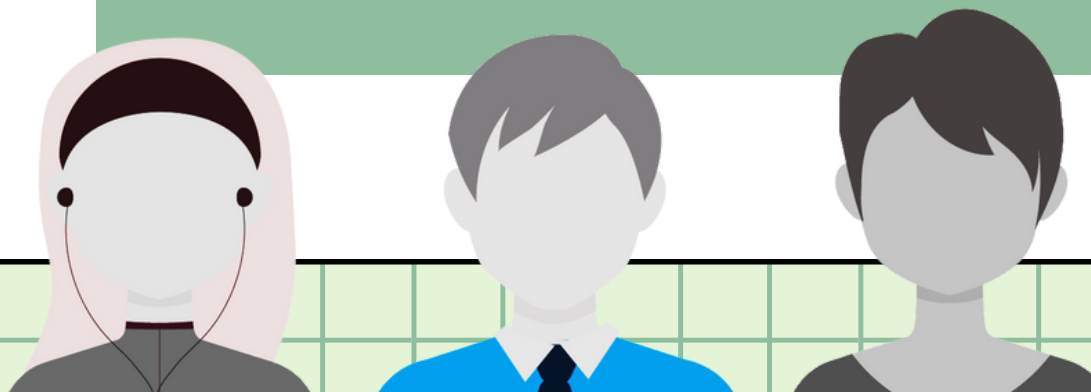
36,5% of all customer

- 542 £ total spend

- 28 orders per year  
= order every 2 weeks/average

- last order 3 weeks ago

> loyal, active customer



# Call to action - how to treat the customer

## UK

- reactivate cluster 2, avoid churn!
- learn about pain-points
- convert basket analysis insights into marketing & promotion
- build product recommendation system

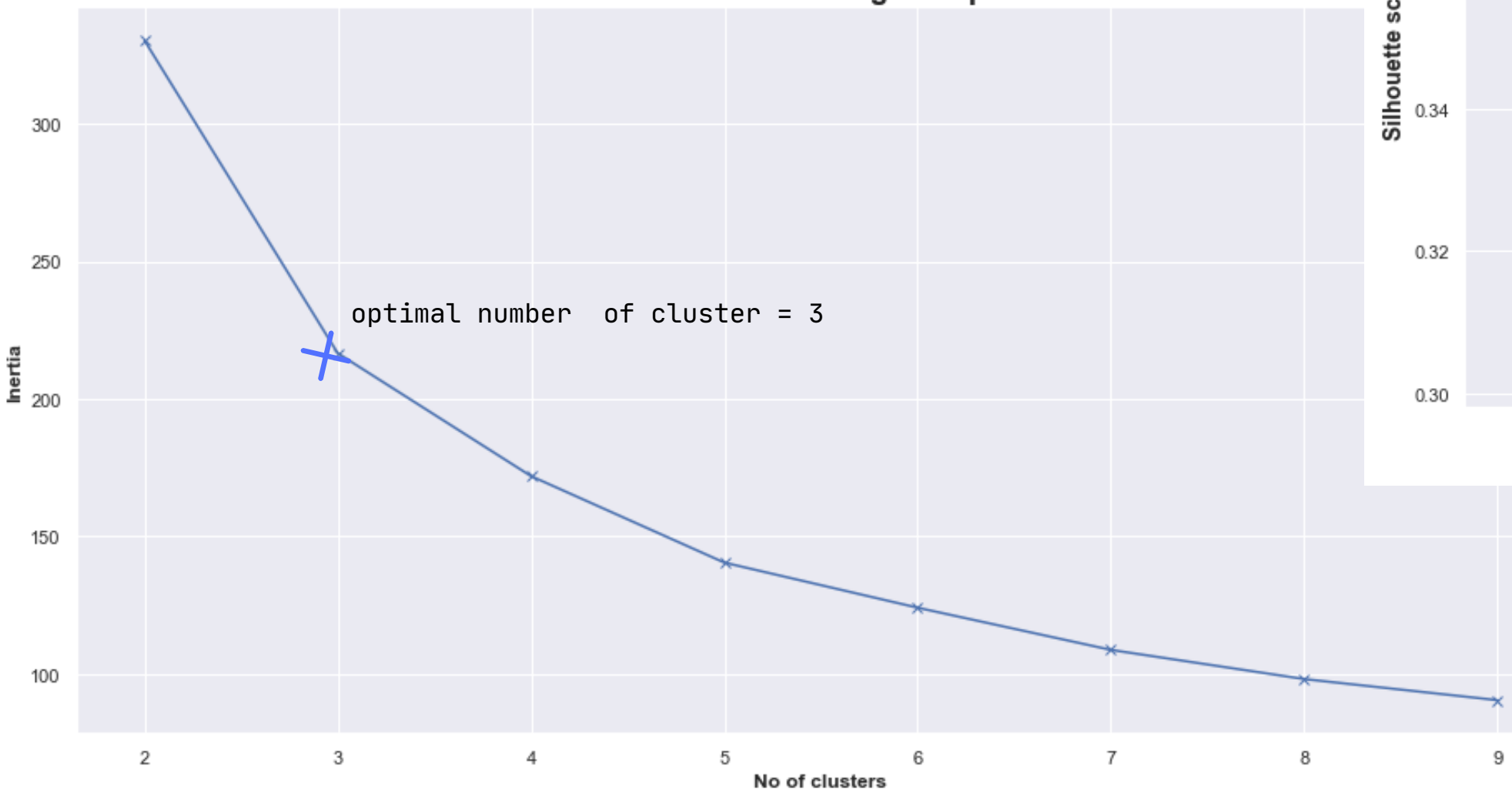


## World

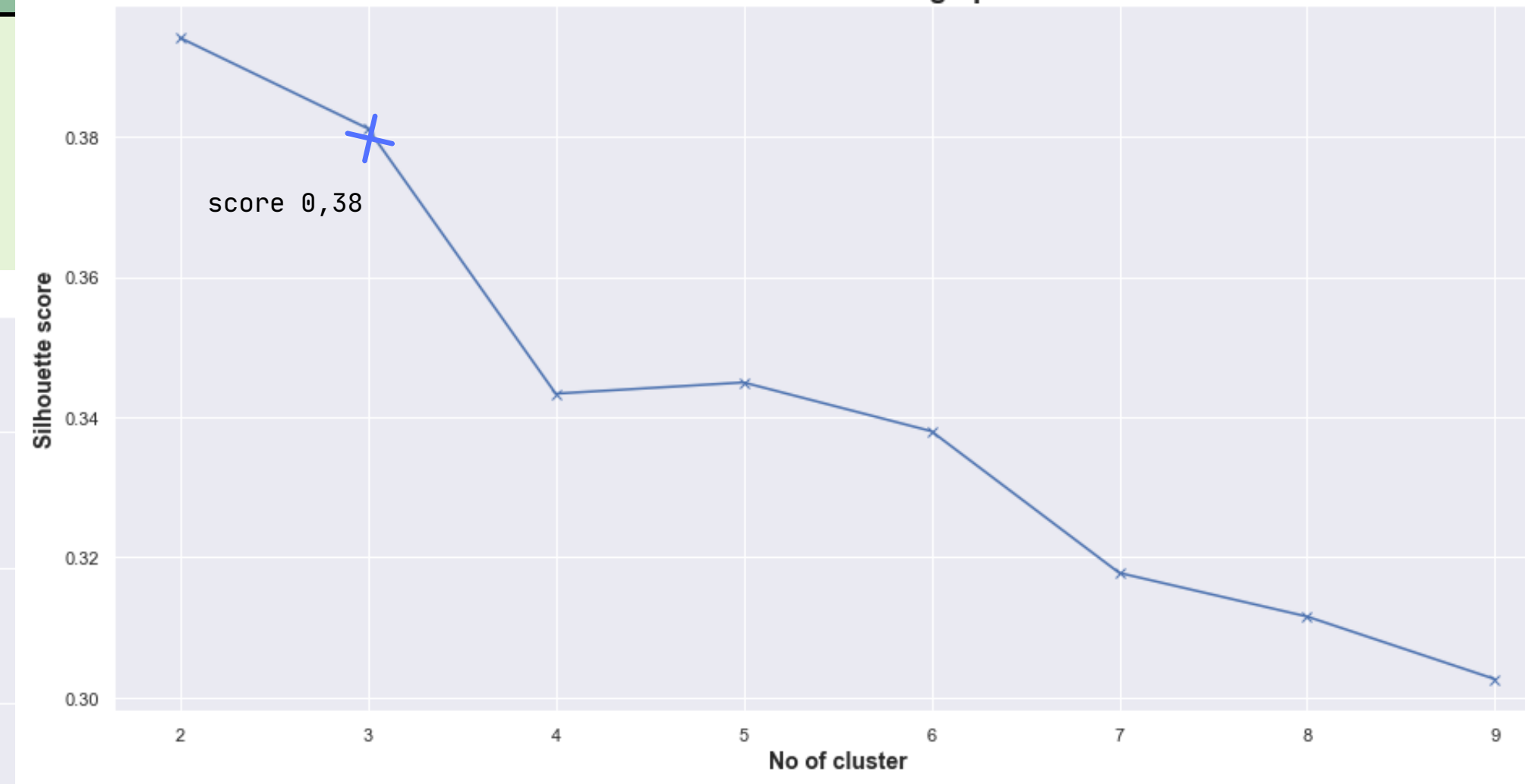
- most valuable customer are international!
- consider region specific marketing campaigns
- Investigate on regional holidays & give special offers
- improve distribution strategy, eg. delivery fee

# K-Means Clustering Workflow

Elbow Method showing the optimal K



Silhouette Curve showing optimal K



These are amazing insights, what else can we do?

### Outlook:

- Deep product analysis & classification of products
- Build recommendation system based on market basket analysis insights with help of clustering
- Lead Scoring with prediction model

**Thank you!**

Want to learn more about customer segmentation  
with algorithms?

I look forward to opportunities to talk about  
that:

Julia Grzejszczak  
jgrzejszczak@outlook.de