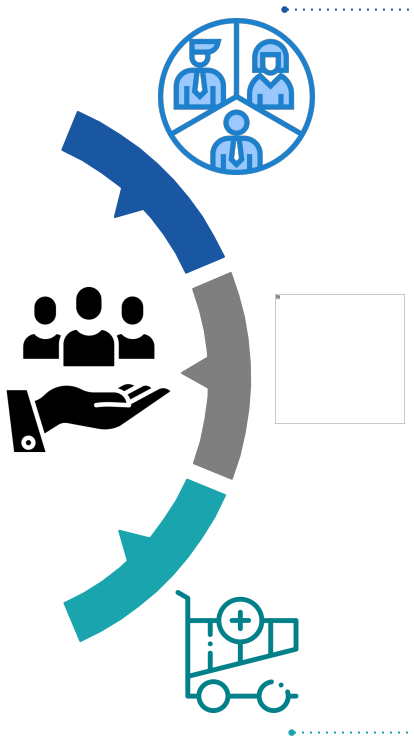
Targeted Marketing Strategy

- Anisha Alluru
- Pooja Dawada
- Satya Pachigolla
- Manas Rai
- Vishal Ramachandran

Overview



Customer Segmentation

- Use RFM and build customer segments to identify focus groups and customize the promotional efforts

Revenue Prediction

- Predict the revenue of a customer to budget expenditure on promotional activities

Recommendation-Market Basket

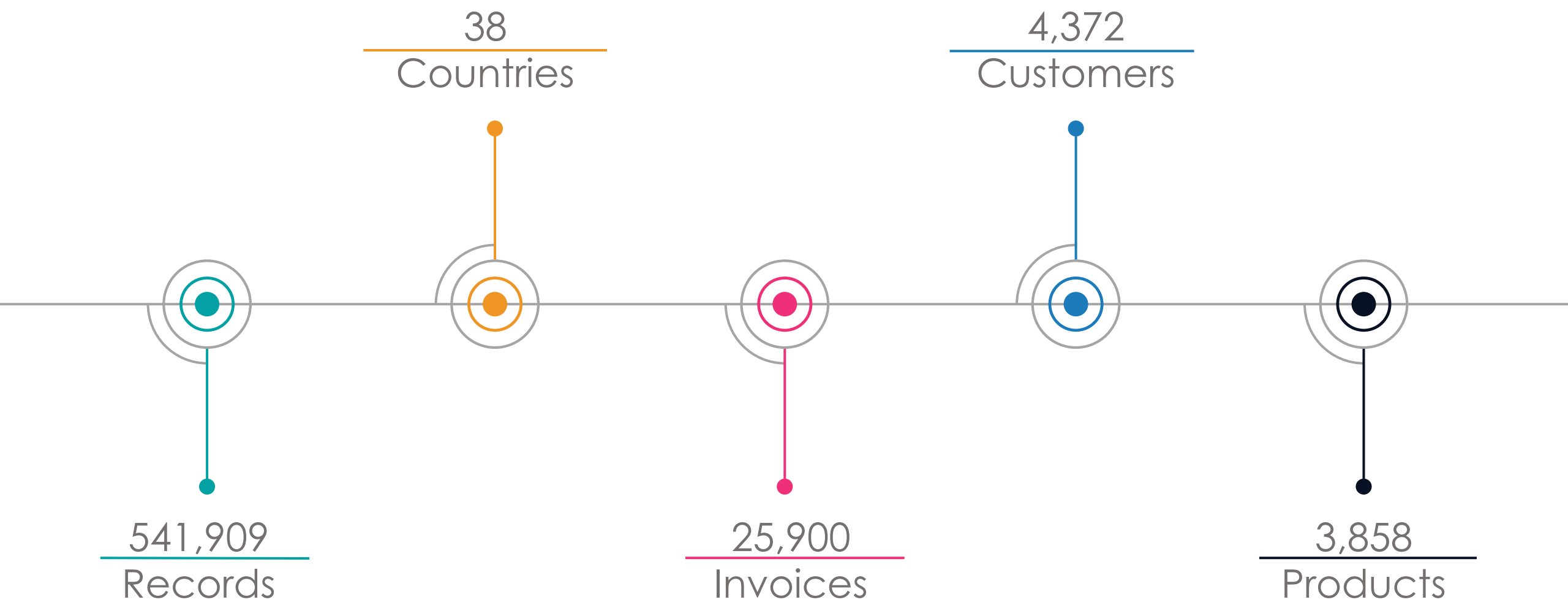
- Identify the categories that we could focus on and design the promotional strategies accordingly

Dataset: https://archive.ics.uci.edu/ml/datasets/online+retail

What does the data look like?

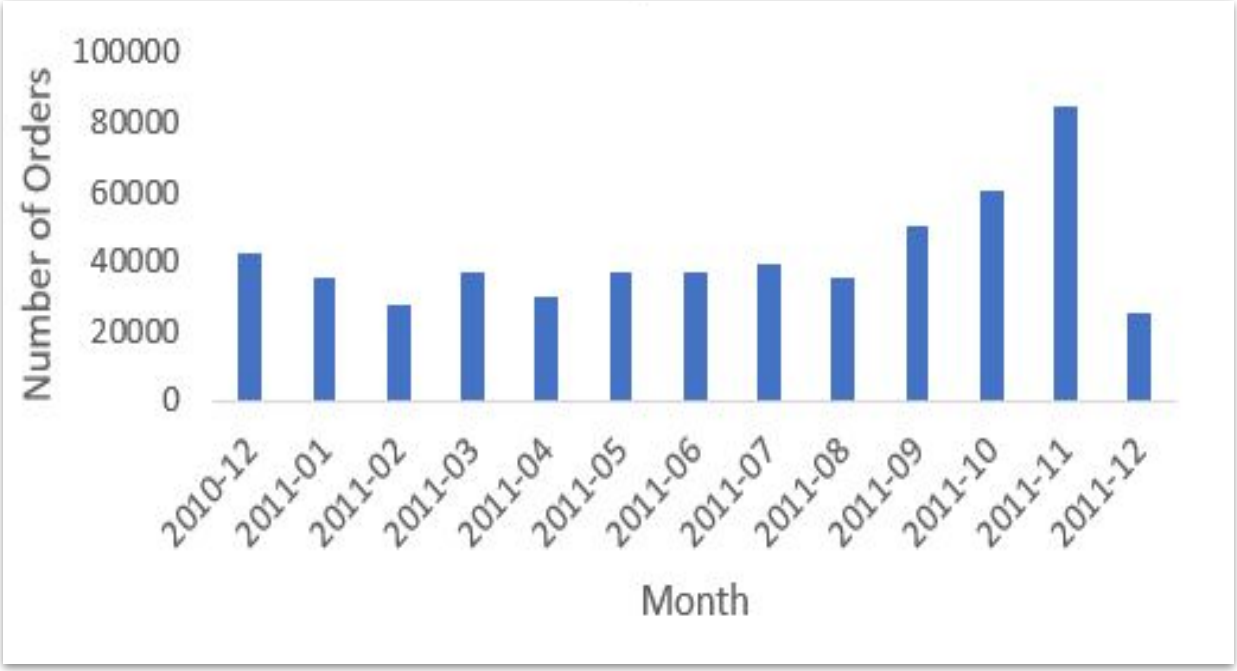
InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850	United Kingdom
536365	71053	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850	United Kingdom
536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	22752	SET 7 BABUSHKA NESTING BOXES	2	12/1/2010 8:26	7.65	17850	United Kingdom
536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	12/1/2010 8:26	4.25	17850	United Kingdom
536366	22633	HAND WARMER UNION JACK	6	12/1/2010 8:28	1.85	17850	United Kingdom
536366	22632	HAND WARMER RED POLKA DOT	6	12/1/2010 8:28	1.85	17850	United Kingdom
536367	84879	ASSORTED COLOUR BIRD ORNAMENT	32	12/1/2010 8:34	1.69	13047	United Kingdom
536367	22745	POPPY'S PLAYHOUSE BEDROOM	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	22748	POPPY'S PLAYHOUSE KITCHEN	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	22749	FELTCRAFT PRINCESS CHARLOTTE DOLL	8	12/1/2010 8:34	3.75	13047	United Kingdom
536367	22310	IVORY KNITTED MUG COSY	6	12/1/2010 8:34	1.65	13047	United Kingdom
536367	84969	BOX OF 6 ASSORTED COLOUR TEASPOONS	6	12/1/2010 8:34	4.25	13047	United Kingdom
536367	22623	BOX OF VINTAGE JIGSAW BLOCKS	3	12/1/2010 8:34	4.95	13047	United Kingdom
536367	22622	BOX OF VINTAGE ALPHABET BLOCKS	2	12/1/2010 8:34	9.95	13047	United Kingdom
536367	21754	HOME BUILDING BLOCK WORD	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	21755	LOVE BUILDING BLOCK WORD	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	21777	RECIPE BOX WITH METAL HEART	4	12/1/2010 8:34	7.95	13047	United Kingdom
536367	48187	DOORMAT NEW ENGLAND	4	12/1/2010 8:34	7.95	13047	United Kingdom
536368	22960	JAM MAKING SET WITH JARS	6	12/1/2010 8:34	4.25	13047	United Kingdom
536368	22913	RED COAT RACK PARIS FASHION	3	12/1/2010 8:34	4.95	13047	United Kingdom

Descriptive Statistics

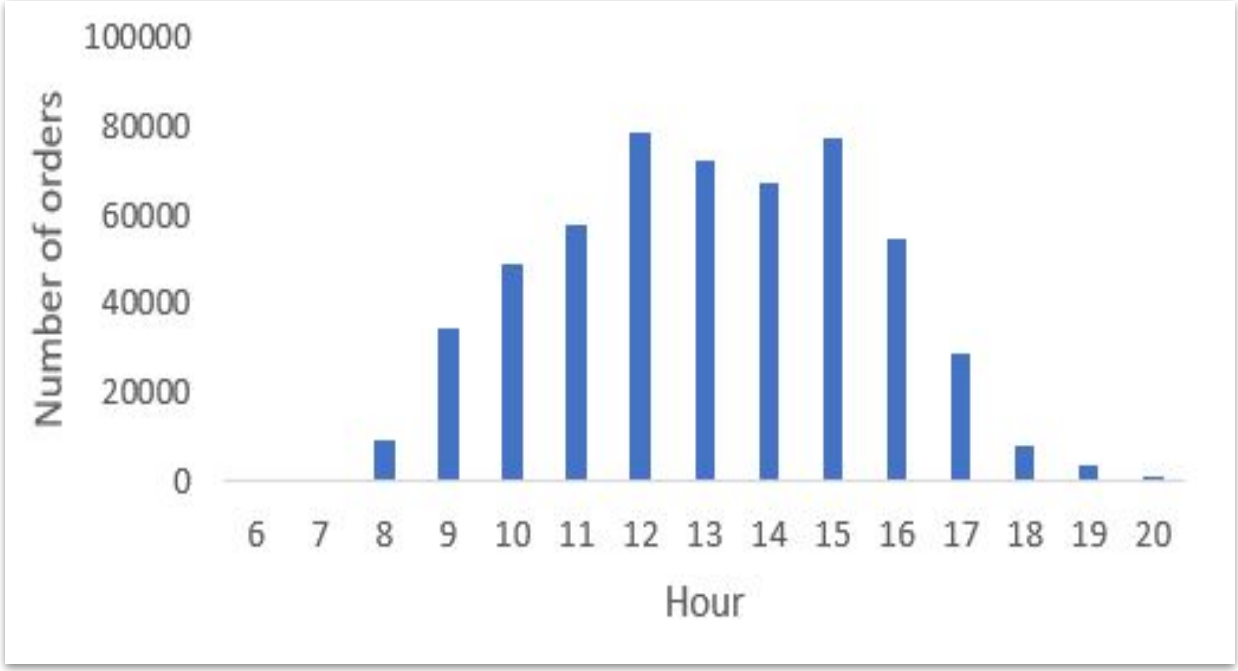


When do people shop?

Orders Per Month



Orders Per Hour

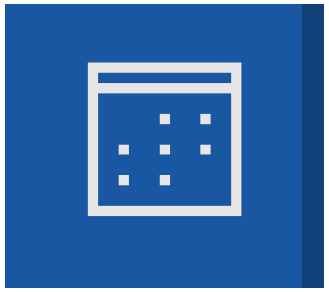


Whom to target?

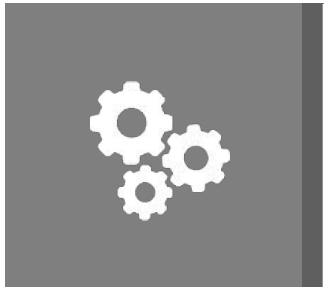
RFM Segmentation


Who did you target?
Everybody
What did it cost?
Everything!!!

Approach



- Considered the first 6 months (Dec'10 – May'11) data to get the RFM segments of the customers



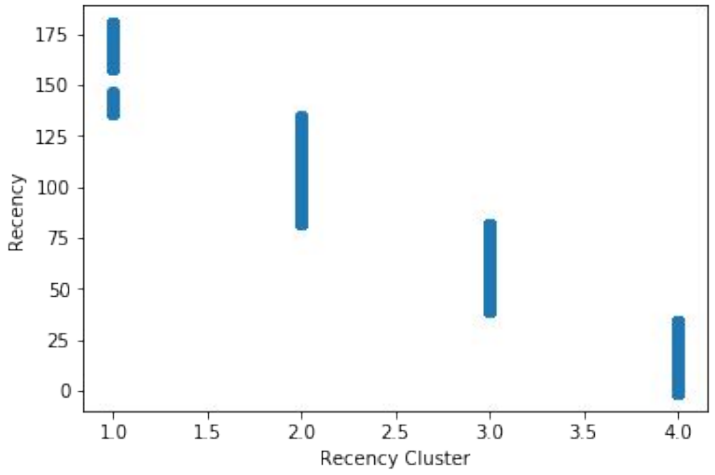
- Removed the transactions which have negative quantities. These are returns
- Removed outliers for the Recency, Frequency and Monetary values
- Did k-means clustering to split the customers into R,F & M clusters

- We will see how the customers move between different RFM clusters in the first and last 6 months
- Using these 6 month RFM clusters we will predict the revenues from the customer for the next 6 months

Recency

- Number of days between the most recent order and the last day (May 31st)
- Divided the customers into 4 clusters

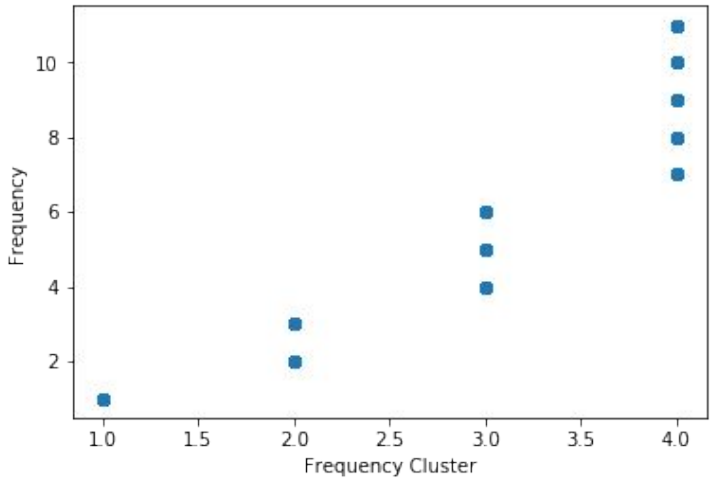
Recency Cluster	# Customers
4	1006
3	717
2	487
1	313



Frequency

- Number of orders made by a customer in the first 6 months
- Divided the customers into 4 clusters

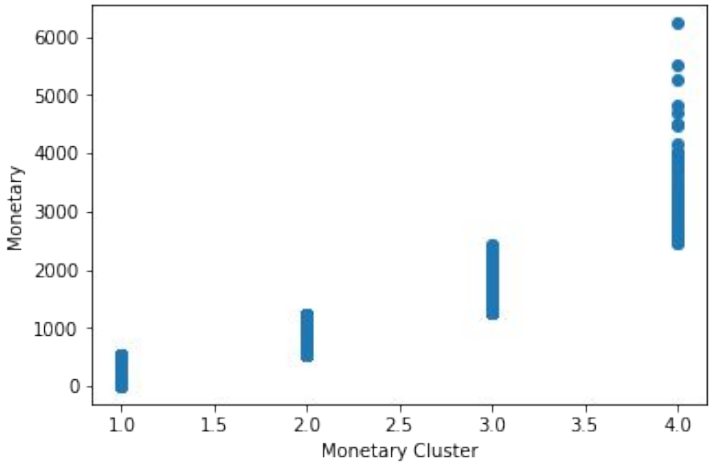
Frequency Cluster	# Customers
4	154
3	418
2	860
1	1091



Monetary

- Revenue from customer in the first 6 months
- Divided the customers into 4 clusters

Monetary Cluster	# Customers
4	89
3	286
2	659
1	1489



Tracking Customer Movements

Customers with medium RFM scores (9,8,7 and 6) are considered ‘Average’

Customers with low RFM scores (5,4 and 3) are considered ‘Low Engagement’

Customers with high RFM scores (12,11 and 10) are considered ‘Loyal’

Customers who are present in first 6 months but did not come back for the next 6 months

% Customers H1	H2 Loyal	Average	Low Engagement	Churn	Grand Total
Loyal	3.80%	4.84%	0.95%	1.51%	11.10%
Average	3.25%	20.93%	12.13%	13.63%	49.94%
Low Engagement	0.79%	10.38%	8.68%	19.10%	38.96%
Grand Total	7.85%	36.15%	21.76%	34.24%	100.00%

How much to spend?

Customer Revenue Prediction

Predicting the next 6 months revenue

1. Used the R,F and M values from the first 6 months for prediction
2. Got an R-squared value of **0.29**

OLS Regression Results

Dep. Variable:	Sales_next6m	R-squared:	0.295
Model:	OLS	Adj. R-squared:	0.294
Method:	Least Squares	F-statistic:	223.1
Date:	Sun, 10 Nov 2019	Prob (F-statistic):	6.58e-121
Time:	15:36:57	Log-Likelihood:	-12735.
No. Observations:	1603	AIC:	2.548e+04
Df Residuals:	1599	BIC:	2.550e+04
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	411.0138	44.438	9.249	0.000	323.851	498.177
Monetory	0.7597	0.040	18.899	0.000	0.681	0.839
Frequency	-9.8132	12.391	-0.792	0.428	-34.117	14.491
Recency	-0.1735	0.383	-0.453	0.651	-0.925	0.578

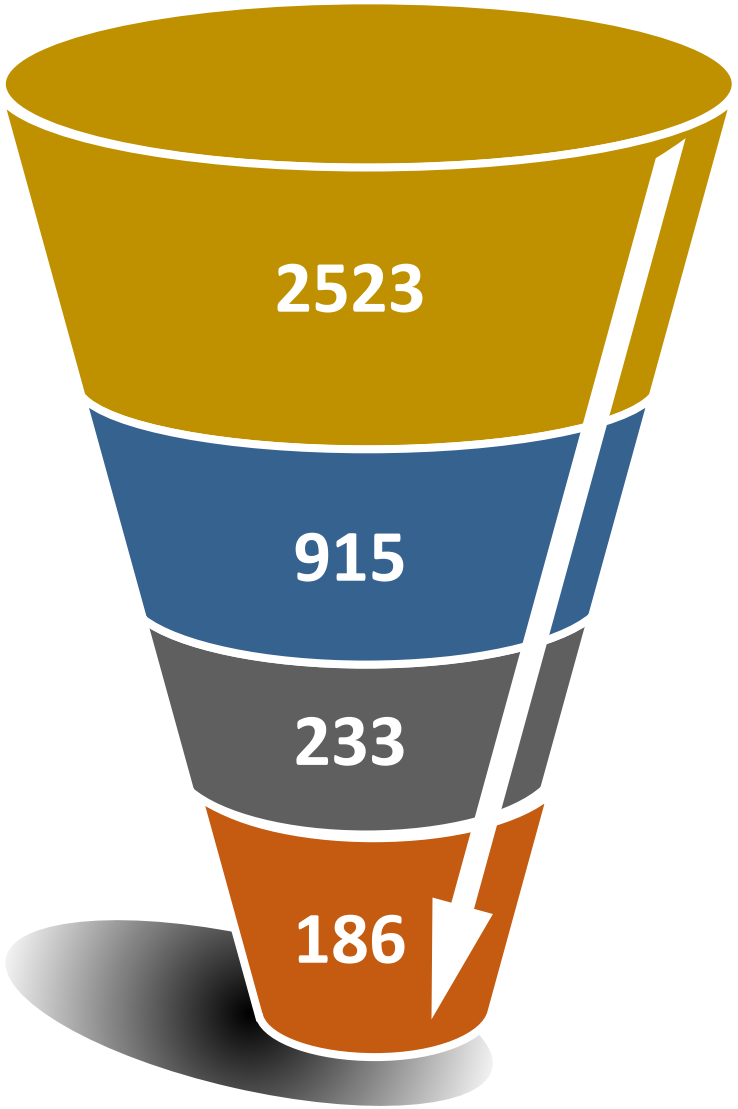
1. Used the R,F and M values from the first 6 months and included the features ‘Active Months’ and ‘Average Basket’
2. Got an R-squared value of **0.77**

OLS Regression Results

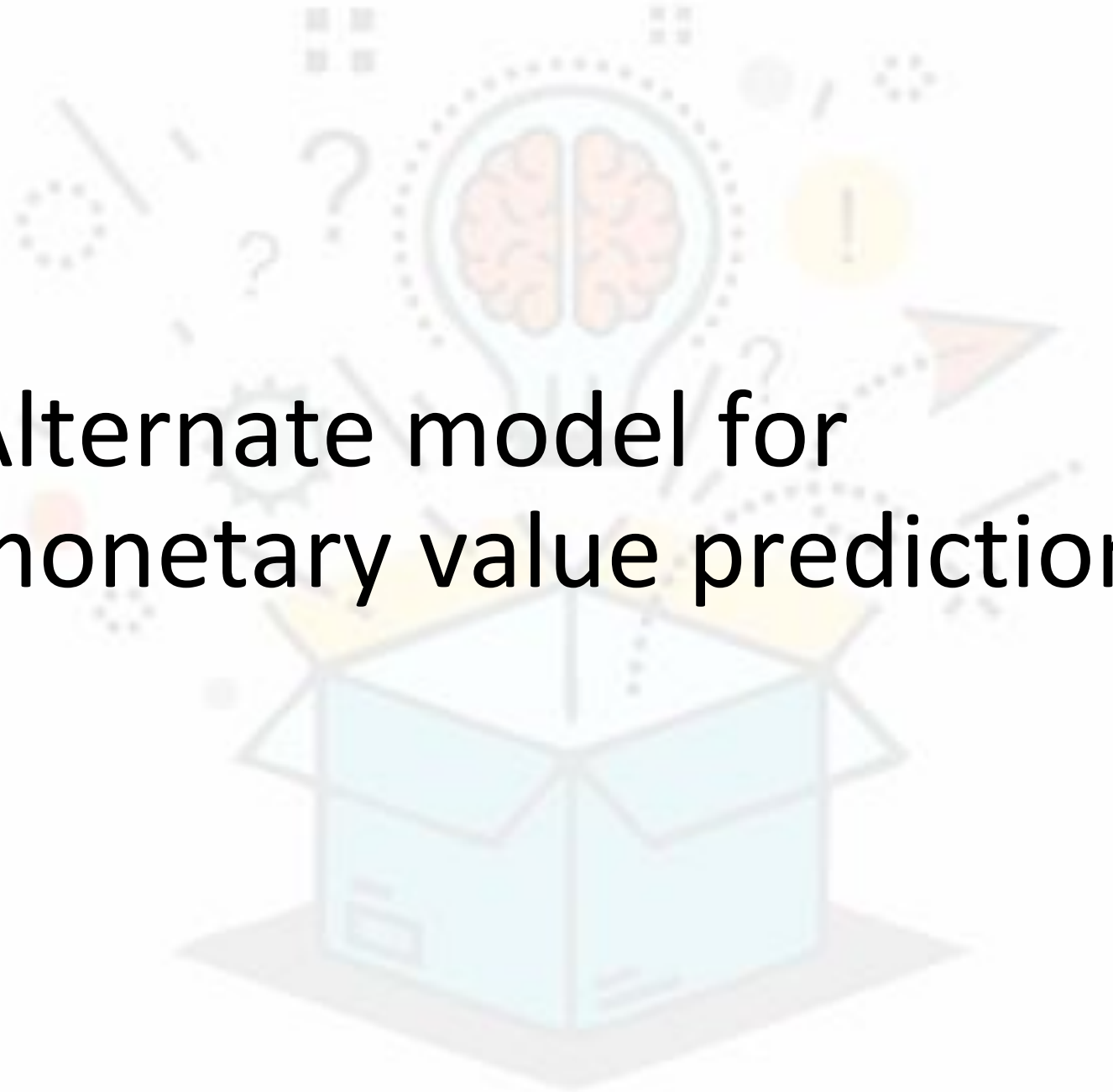
Dep. Variable:	Sales_next6m	R-squared:	0.775
Model:	OLS	Adj. R-squared:	0.774
Method:	Least Squares	F-statistic:	1100.
Date:	Sun, 10 Nov 2019	Prob (F-statistic):	0.00
Time:	15:34:53	Log-Likelihood:	-11820.
No. Observations:	1603	AIC:	2.365e+04
Df Residuals:	1597	BIC:	2.368e+04
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-784.7506	33.080	-23.723	0.000	-849.635	-719.866
Monetory	0.2318	0.026	9.009	0.000	0.181	0.282
Frequency	-14.2428	7.498	-1.899	0.058	-28.950	0.465
Recency	0.4570	0.218	2.100	0.036	0.030	0.884
ActiveMonths	408.4792	8.062	50.665	0.000	392.665	424.293
Average Basket	2.0853	0.066	31.756	0.000	1.956	2.214

Focused Targeting

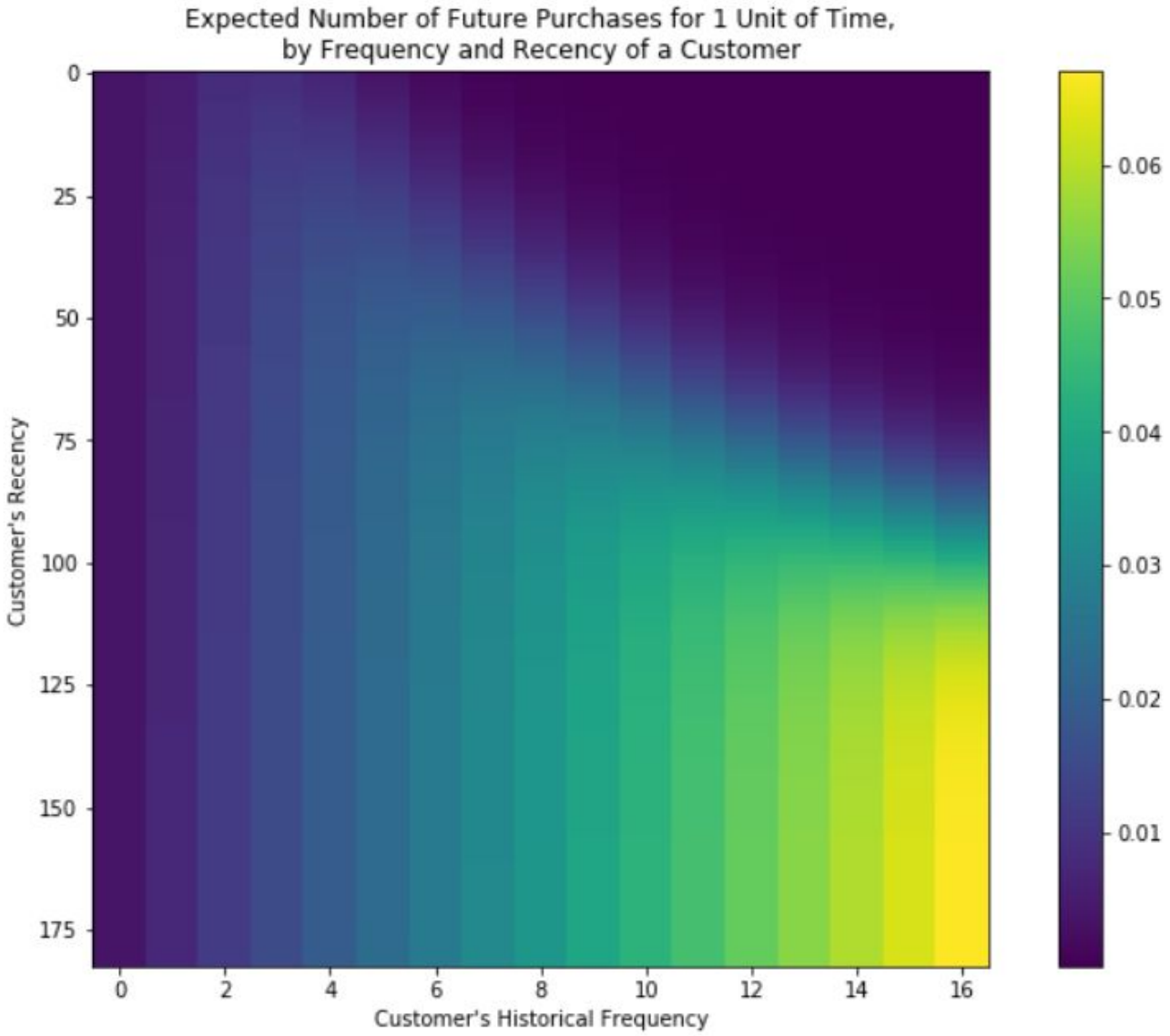


- Customers who shopped in the first 6 months
- Customers whom we want to target – the **‘Average’** base
- Customers whose ‘**Predicted**’ sales decline is greater than 30%
- Among the 233, customers whose actual sales decline is greater than 30%



Alternate model for monetary value prediction

Beta Geo Fitter – Purchase Frequency



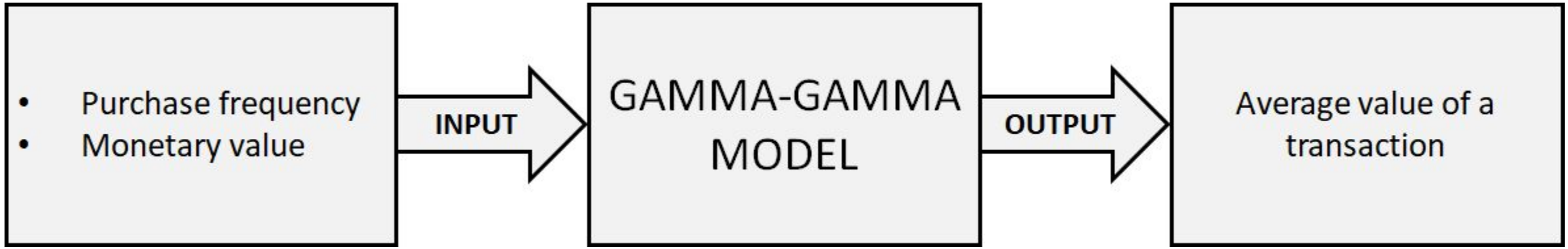
Input

- No. of days active
- Current purchase frequency

Output

- Frequency of purchase per unit time in the future

Gamma-Gamma Model – Mean spend per purchase



- The mean predicted by this model will be the mean value of all transactions in a customer's lifetime
- The value of each transaction in the future adds random noise around this mean value

Sample Calculation

CustomerID - 17389

Frequency - 14

Recency - 140

Beta Geo frequency/per unit time = 0.07

No. of purchases in next 6 months = 0.07*180 = 12.6

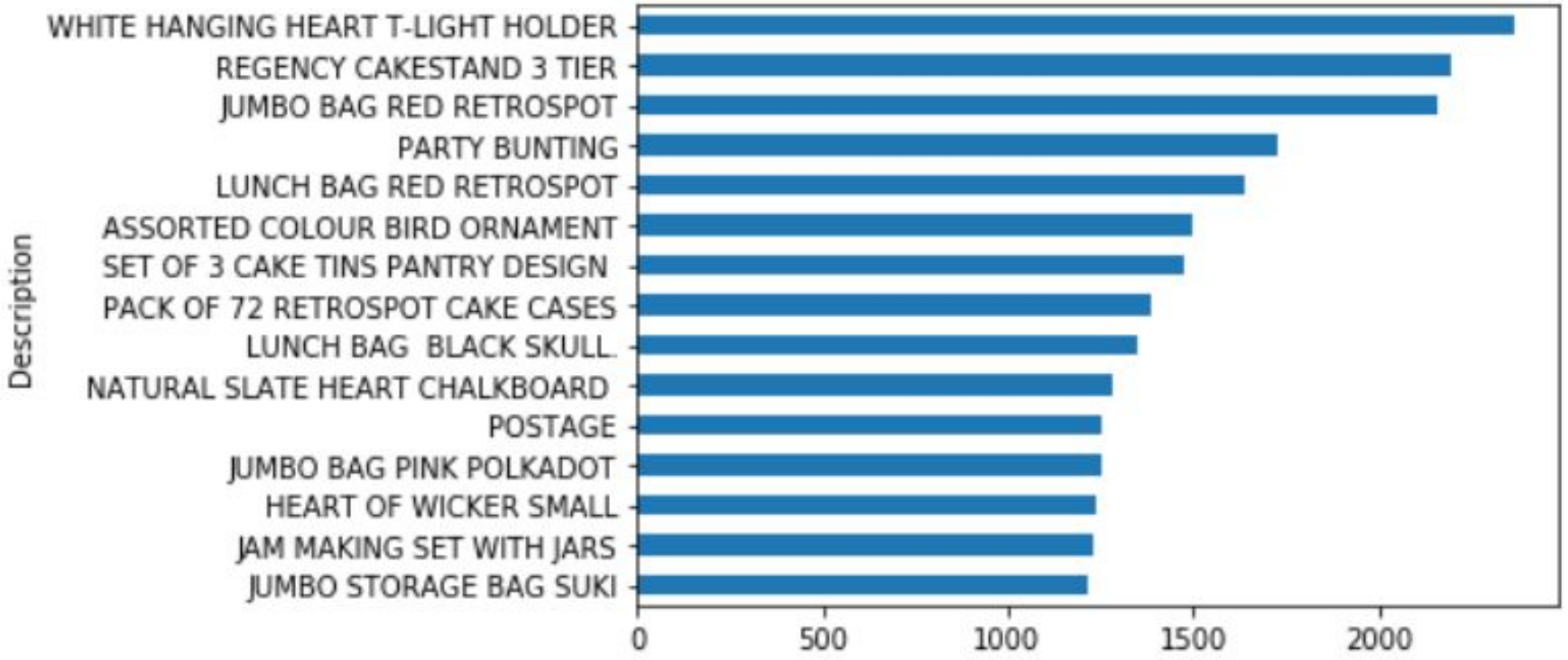
Gamma-Gamma average basket value = 353.04

Monetary value of the customer in next months = 353.04*12.6 = $4448.3

What to target them with?

Market Basket Analysis

Hot Sellers


Description
WHITE HANGING HEART T-LIGHT HOLDER
REGENCY CAKESTAND 3 TIER
JUMBO BAG RED RETROSPOT
PARTY BUNTING
LUNCH BAG RED RETROSPOT
ASSORTED COLOUR BIRD ORNAMENT
SET OF 3 CAKE TINS PANTRY DESIGN
PACK OF 72 RETROSPOT CAKE CASES
LUNCH BAG BLACK SKULL.
NATURAL SLATE HEART CHALKBOARD
POSTAGE
JUMBO BAG PINK POLKADOT
HEART OF WICKER SMALL
JAM MAKING SET WITH JARS
JUMBO STORAGE BAG SUKI
0
500
1000
1500
2000

Market Basket Analysis

$$Rule: X \Rightarrow Y$$

$$Support = \frac{frq(X,Y)}{N}$$

$$Confidence = \frac{frq(X,Y)}{frq(X)}$$

$$Lift = \frac{Support}{Supp(X) \times Supp(Y)}$$

Rule	Support	Confidence	Lift
$A \Rightarrow D$	2/5	2/3	10/9
$C \Rightarrow A$	2/5	2/4	5/6
$A \Rightarrow C$	2/5	2/3	5/6
$B \& C \Rightarrow D$	1/5	1/3	5/9

Summary of the rules

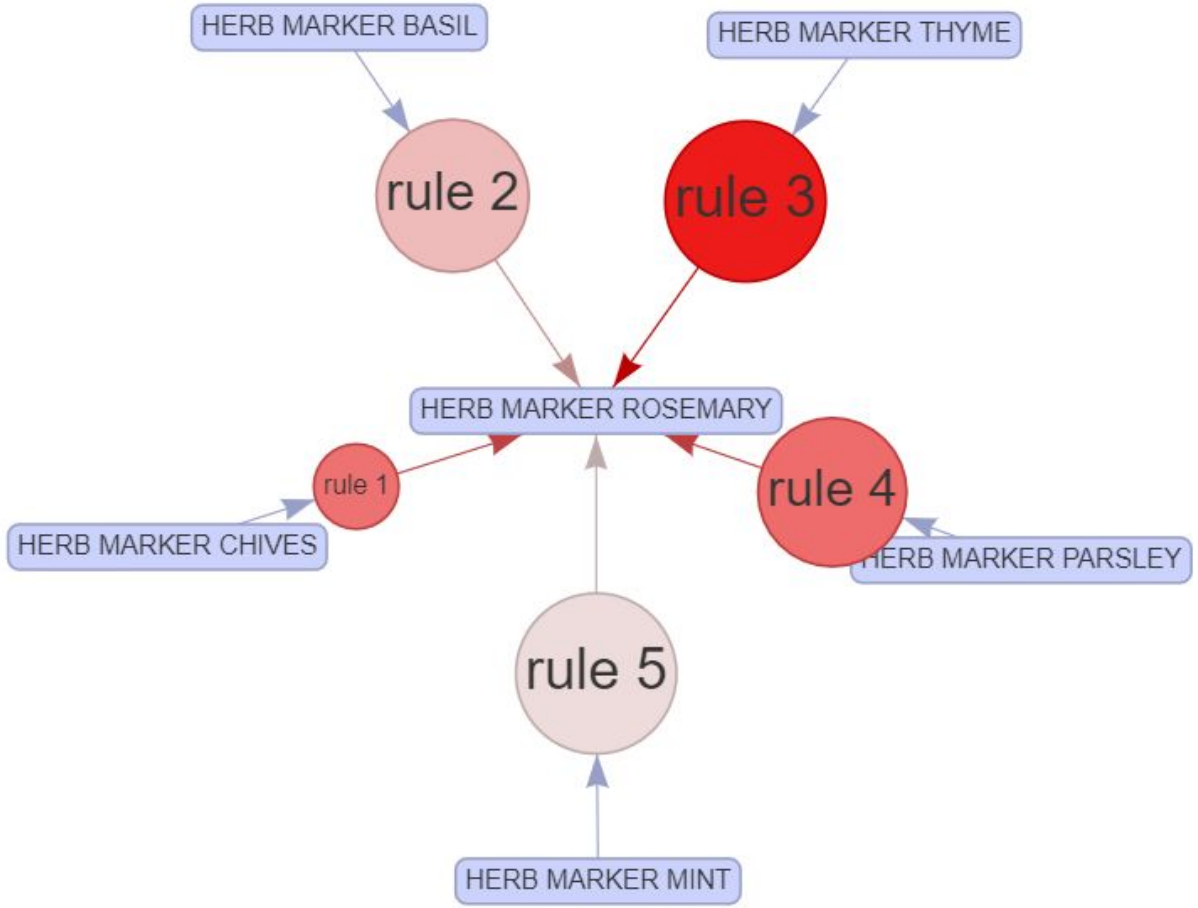
~1400 rules

- Support threshold = 0.01
- Confidence threshold = 0.70
- Lift threshold - 10

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift
1941	(HERB MARKER PARSLEY, HERB MARKER BASIL)	(HERB MARKER THYME, HERB MARKER MINT, HERB MAR...	0.014182	0.013818	0.010909	0.769231	55.668016
196	(COFFEE MUG DOG + BALL DESIGN)	(COFFEE MUG CAT + BIRD DESIGN)	0.012727	0.015273	0.012364	0.971429	63.605442
1244	(HERB MARKER PARSLEY)	(HERB MARKER THYME, HERB MARKER ROSEMARY)	0.016727	0.014545	0.013455	0.804348	55.298913
831	(RECYCLED ACAPULCO MAT GREEN)	(RECYCLED ACAPULCO MAT TURQUOISE)	0.015273	0.021455	0.011273	0.738095	34.402744
1230	(HERB MARKER MINT, HERB MARKER PARSLEY)	(HERB MARKER THYME)	0.014909	0.016364	0.013455	0.902439	55.149051
988	(CHILDS GARDEN FORK BLUE, CHILDS GARDEN TROWEL...	(CHILDS GARDEN FORK PINK)	0.011636	0.014182	0.009818	0.843750	59.495192
93	(KITCHEN METAL SIGN)	(BATHROOM METAL SIGN)	0.013091	0.025818	0.010182	0.777778	30.125196
559	(KEY FOB , GARAGE DESIGN)	(KEY FOB , BACK DOOR)	0.013818	0.013091	0.010545	0.763158	58.296784
1651	(HERB MARKER ROSEMARY)	(HERB MARKER THYME, HERB MARKER MINT, HERB MAR...	0.016000	0.012364	0.012000	0.750000	60.661765
1529	(HERB MARKER MINT, HERB MARKER BASIL, HERB MAR...	(HERB MARKER PARSLEY)	0.010182	0.016727	0.010182	1.000000	59.782609
904	(SET/6 RED SPOTTY PAPER PLATES)	(SET/6 RED SPOTTY PAPER CUPS)	0.020364	0.018182	0.017091	0.839286	46.160714

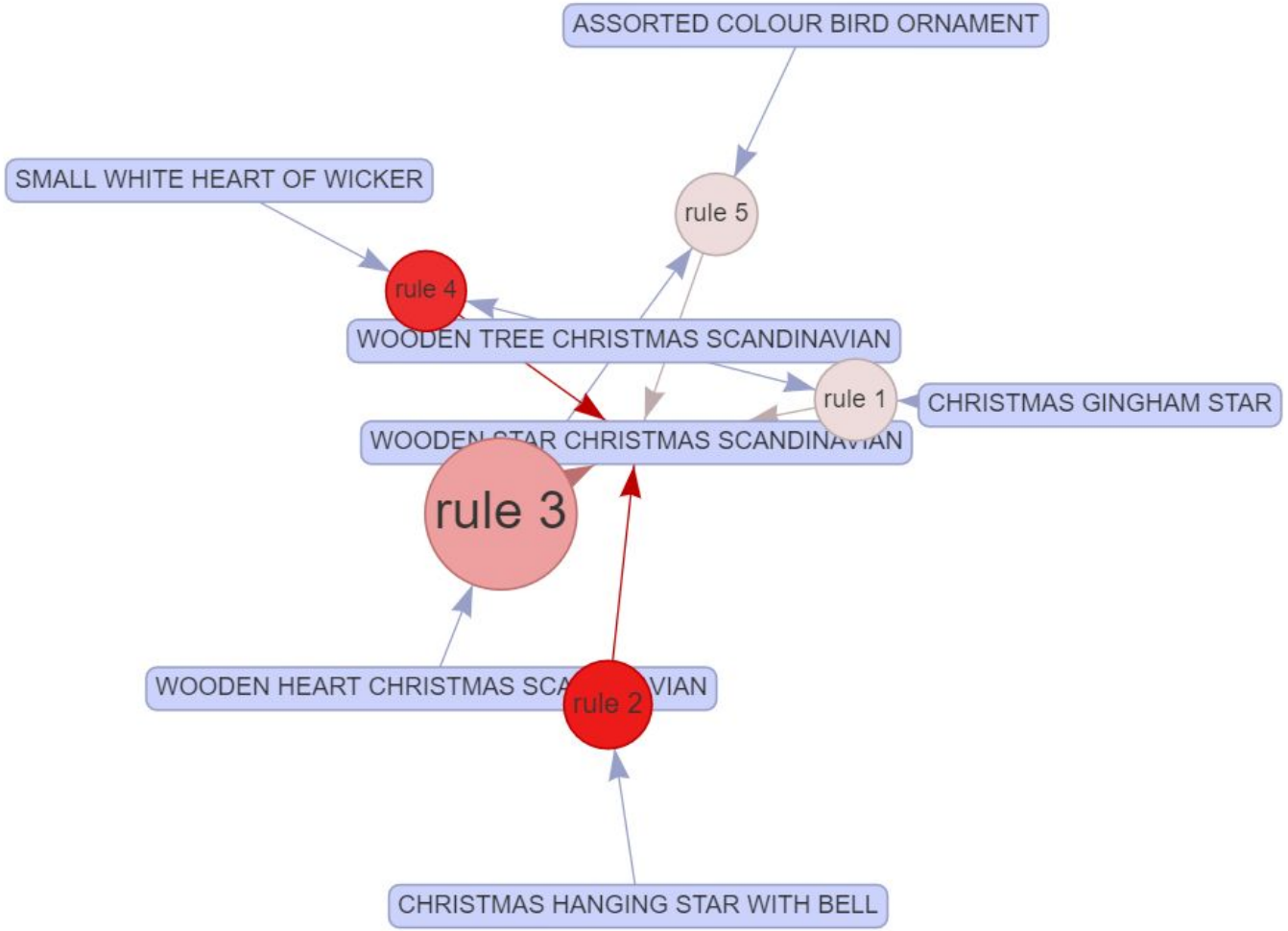
Most Frequently Bought Together...

Support	1%
Confidence	93%
Lift	91



Most Frequently Bought Together...

Support	1%
Confidence	91%
Lift	39



Most Frequently Bought Together...

Executive Summary

We need better marketing strategy to retain customers falling in the “Average” bucket

Targeting the customers we identify to be dropping by 30% in sales, would maximize the benefits

We can offer them highly affined products as bundles to attract them back to the site


IT'S OVER
IT'S FINALLY OVER
quickmeme.com