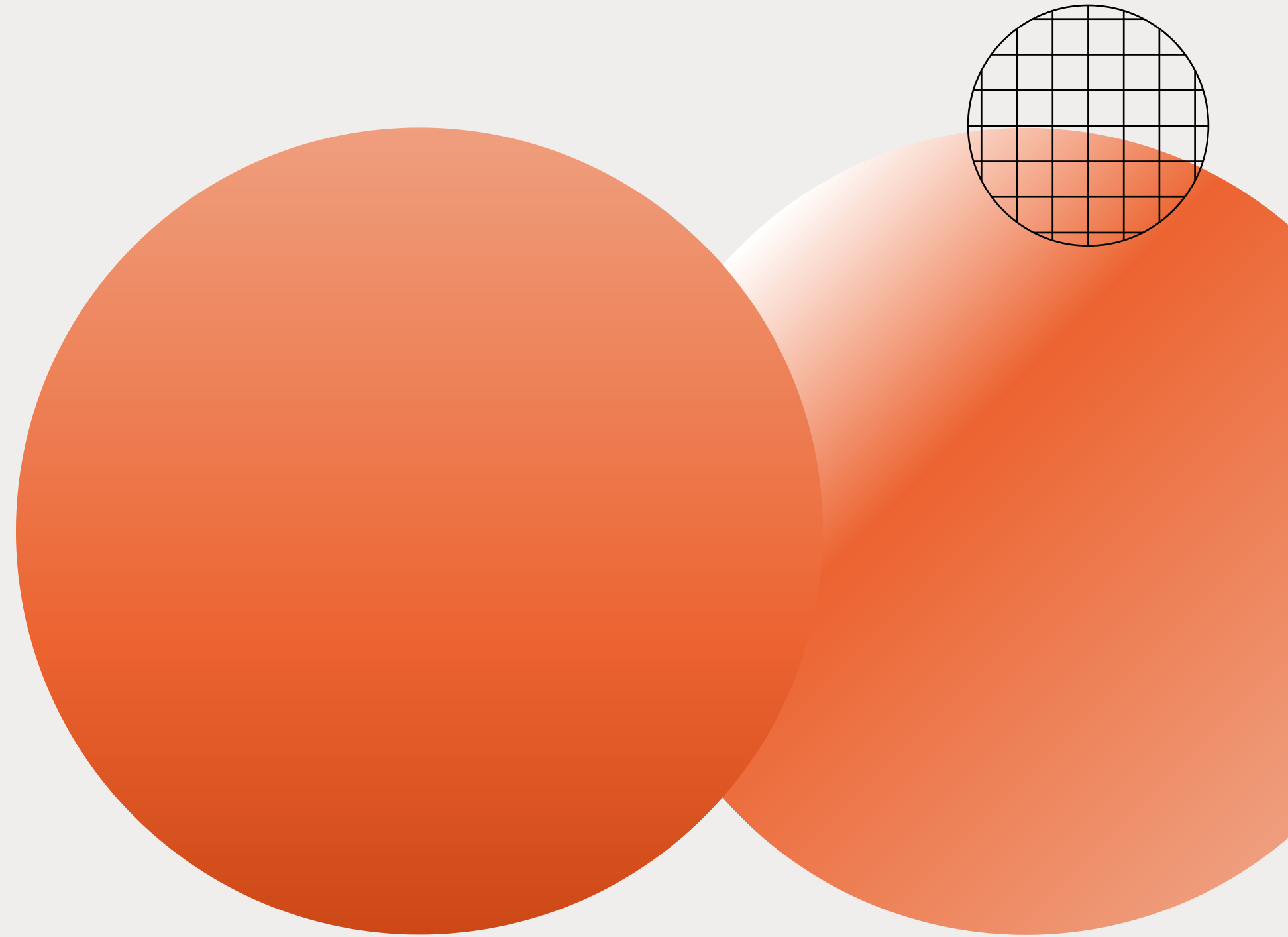


Gotta catch 'em all!

Targeted Advertising on
identified key segments to
maximize customer retention



Overview

- 1 Customer Segmentation
- 2 Purchase Probability Prediction
- 3 Recommendation

Scope of Project

Identify focus groups using RFM analysis to understand which customer segments are most likely to be swayed by Targeted Advertising to maximize customer growth and retention.



Dataset Sneak Peek

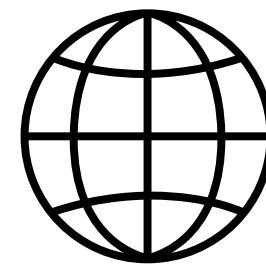
InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
536365	85123A	WHITE HANGING HE	6	12/1/2010 8:26	2.55	17850	United Kingdom
536365	71053	WHITE METAL LANT	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	84406B	CREAM CUPID HEAR	8	12/1/2010 8:26	2.75	17850	United Kingdom
536365	84029G	KNITTED UNION FLA	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	84029E	RED WOOLLY HOTTI	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	22752	SET 7 BABUSHKA NE	2	12/1/2010 8:26	7.65	17850	United Kingdom
536365	21730	GLASS STAR FROSTE	6	12/1/2010 8:26	4.25	17850	United Kingdom
536366	22633	HAND WARMER UN	6	12/1/2010 8:28	1.85	17850	United Kingdom
536366	22632	HAND WARMER REC	6	12/1/2010 8:28	1.85	17850	United Kingdom
536367	84879	ASSORTED COLOUR	32	12/1/2010 8:34	1.69	13047	United Kingdom
536367	22745	POPPY'S PLAYHOUSE	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	22748	POPPY'S PLAYHOUSE	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	22749	FELTCRAFT PRINCES	8	12/1/2010 8:34	3.75	13047	United Kingdom
536367	22310	IVORY KNITTED MU	6	12/1/2010 8:34	1.65	13047	United Kingdom
536367	84969	BOX OF 6 ASSORTED	6	12/1/2010 8:34	4.25	13047	United Kingdom
536367	22623	BOX OF VINTAGE JIG	3	12/1/2010 8:34	4.95	13047	United Kingdom
536367	22622	BOX OF VINTAGE AL	2	12/1/2010 8:34	9.95	13047	United Kingdom
536367	21754	HOME BUILDING BL	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	21755	LOVE BUILDING BLC	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	21777	RECIPE BOX WITH M	4	12/1/2010 8:34	7.95	13047	United Kingdom
536367	48187	DOORMAT NEW EN	4	12/1/2010 8:34	7.95	13047	United Kingdom
536368	22960	JAM MAKING SET W	6	12/1/2010 8:34	4.25	13047	United Kingdom
536368	22913	RED COAT RACK PAR	3	12/1/2010 8:34	4.95	13047	United Kingdom
536368	22912	YELLOW COAT RACK	3	12/1/2010 8:34	4.95	13047	United Kingdom
536368	22914	BLUE COAT RACK PA	3	12/1/2010 8:34	4.95	13047	United Kingdom
536369	21756	BATH BUILDING BLC	3	12/1/2010 8:35	5.95	13047	United Kingdom

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

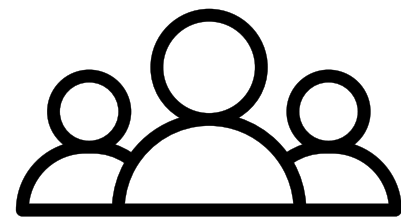
Dataset Highlights



1) Number of Instances:
541909



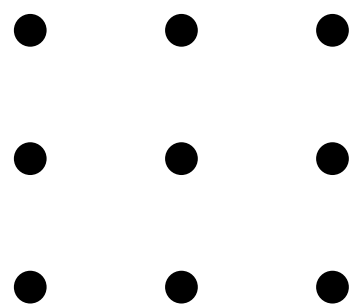
2) Number of Attributes:
8



3) Time Frame:
2010-12-01 to 2011-12-
09



4) Number of Customers:
4372



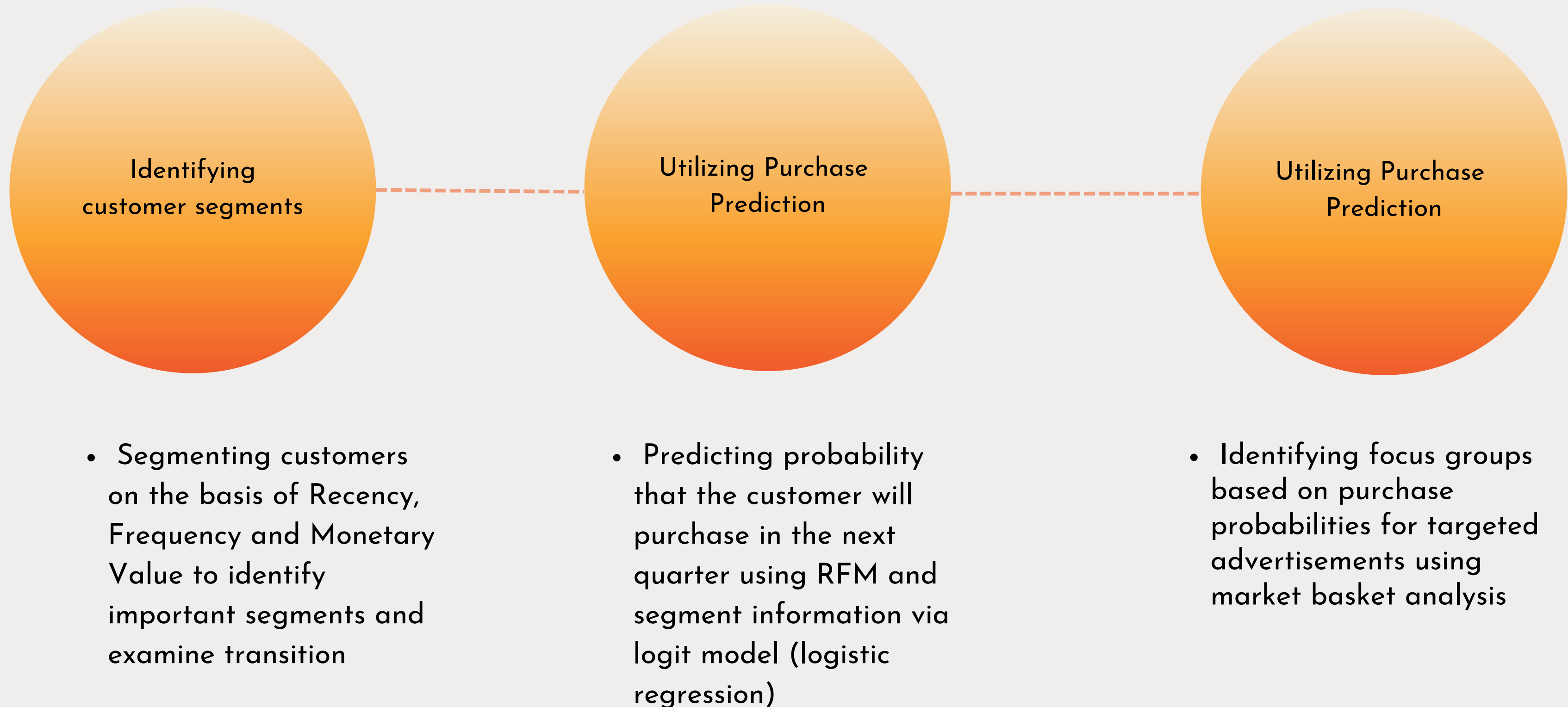
Data Preprocessing

- 1) Limited Country scope to only the United Kingdom
- 2) Dropped all null CustomerID values for RFM segmentation

Result: Dataset size reduced to 73% of the original and # of Customers reduced to 3950

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

Approach



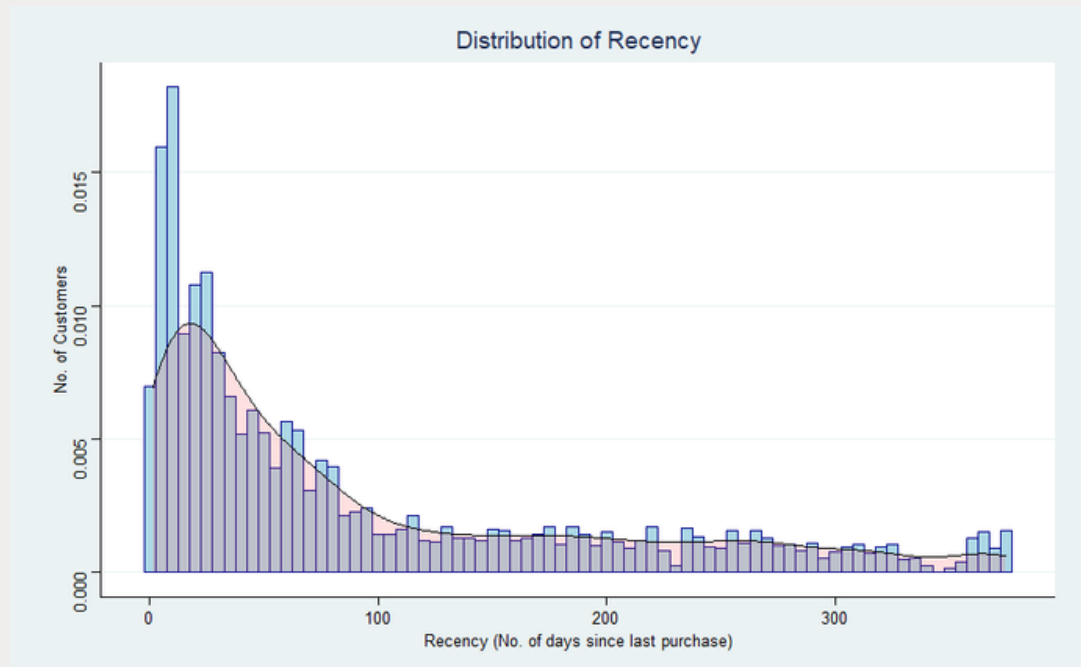


Customer Segmentation

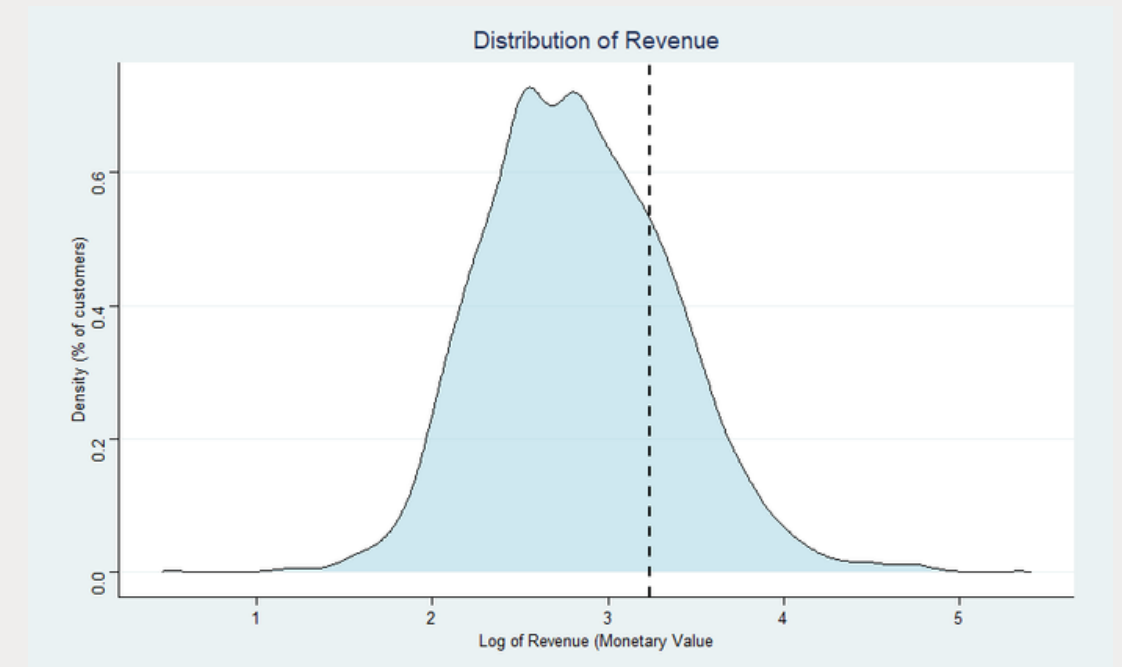
RFM Analysis



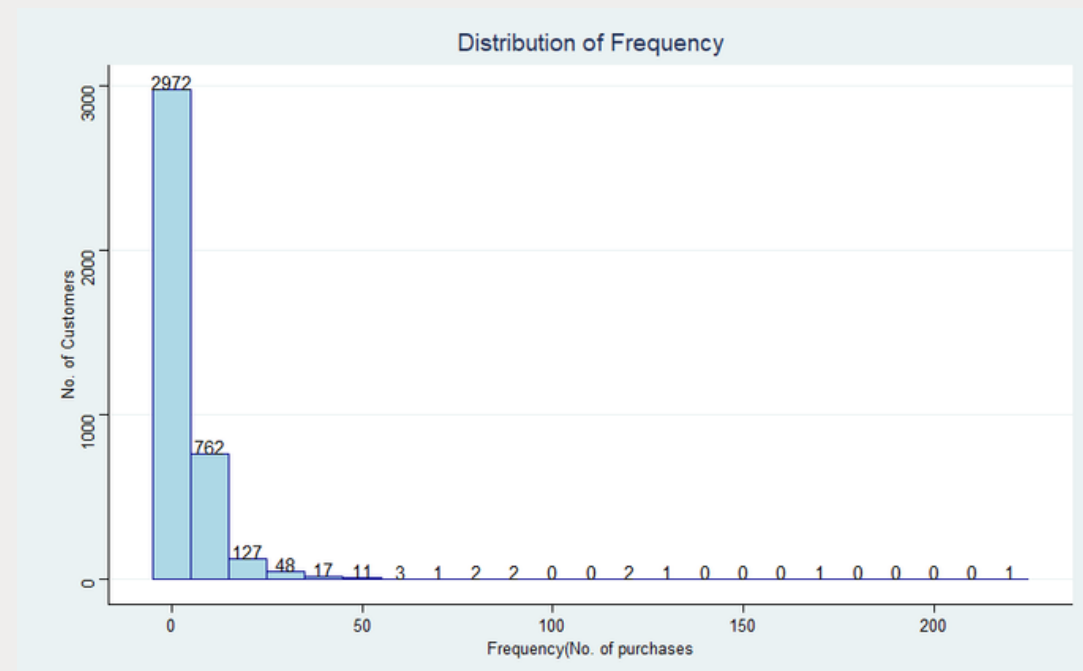
Segmentation Thresholds



- **Guest** - Between 0 to 5 purchases (median)
- **Recurring** - More than 5 purchases

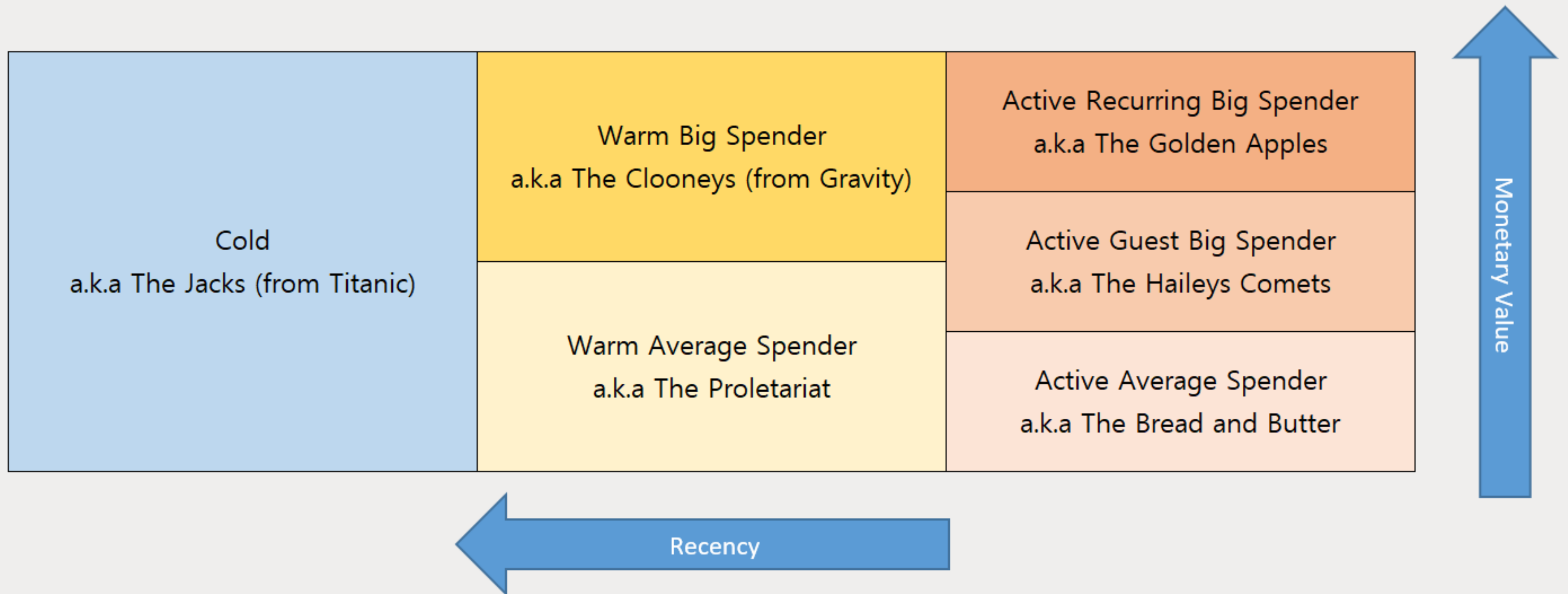


- **Active** - Between 0 days and 51 days (0 to 1st Quartile)
- **Warm** - Between 51 days and 144 days (1st to 3rd Quartile)
- **Cold** - More than 144 days old (beyond 3rd Quartile)

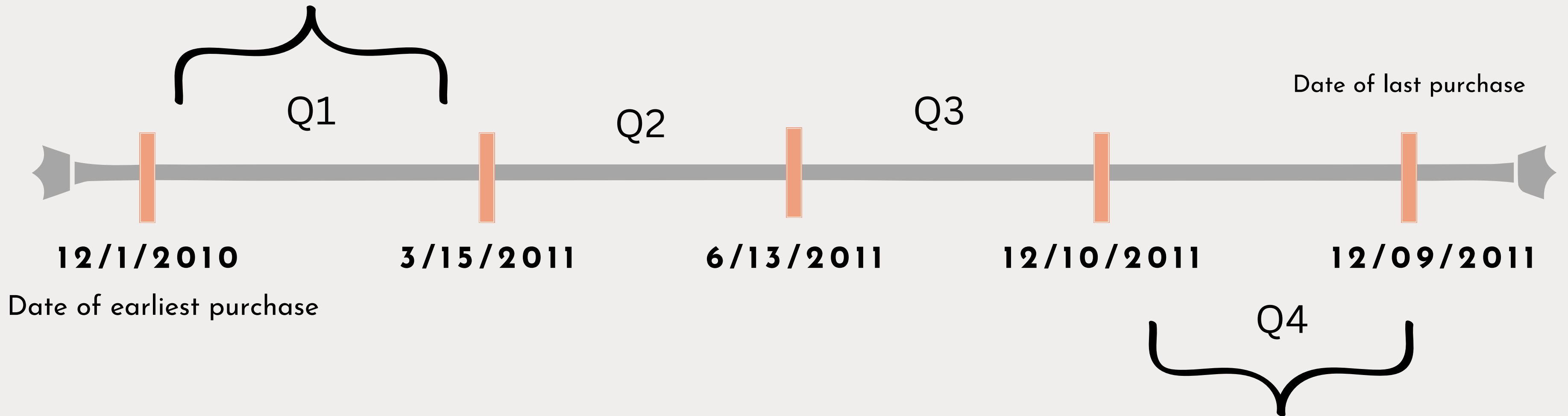


- **Big Spender** - More than 1713 pounds spent in total (mean)
- **Average Spender** - Less than or equal to 1713 pounds spent in total

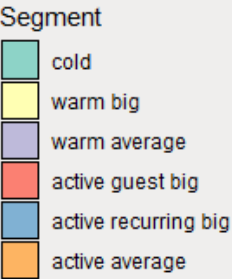
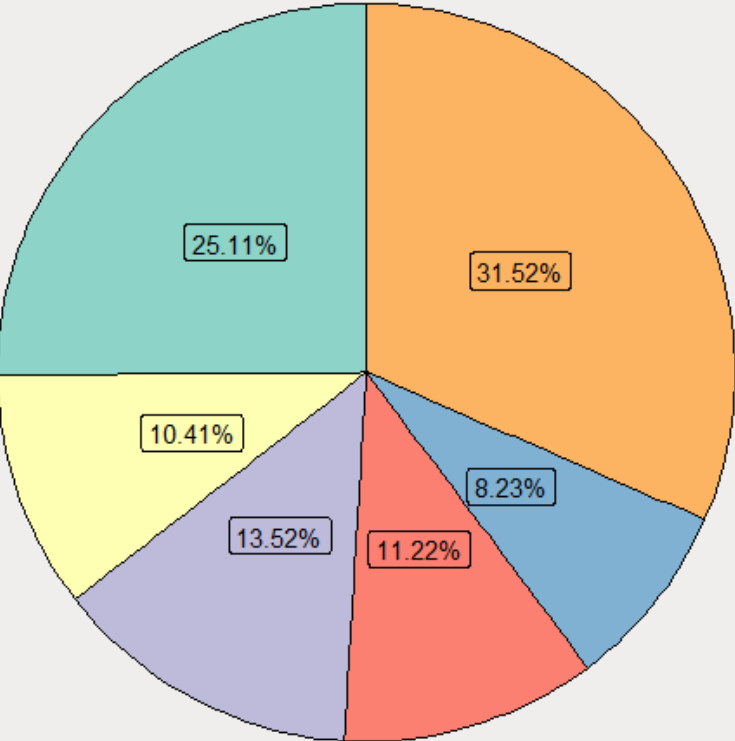
Customer Segments



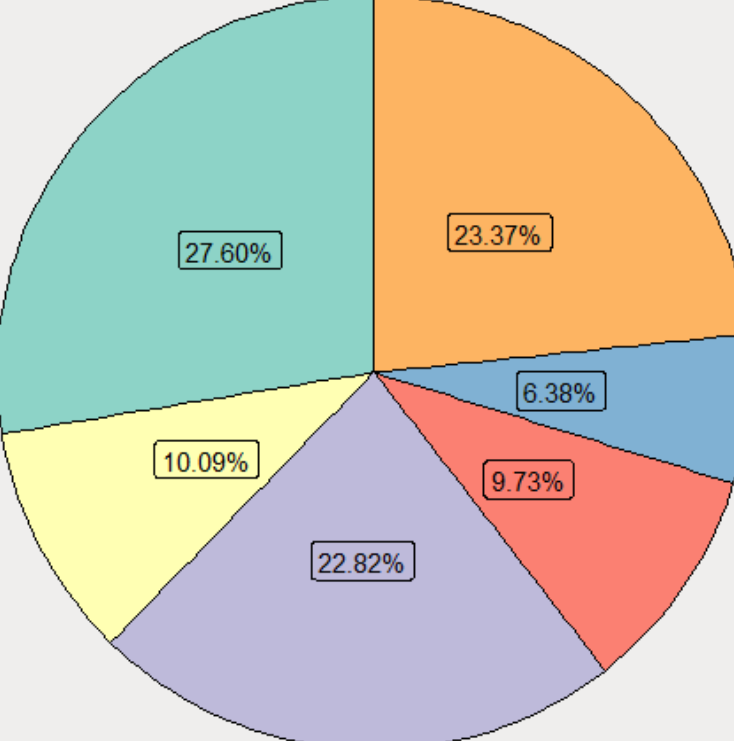
Transition Timeline



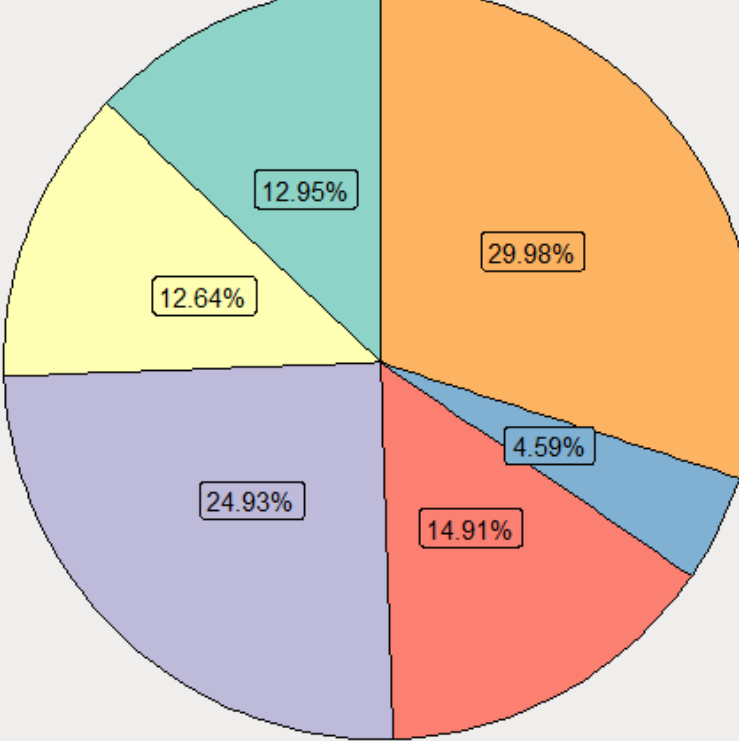
Composition



QUARTER
2



QUARTER
3



QUARTER
4

Matrix of Transitions

QUARTER 2

QUARTER 3							
	Cold	Warm Average Spender	Warm Big Spender	Active Average Spender	Active Guest Big Spender	Active Recurring Big Spender	Total
Cold	249	32	9	26	19	1	336
Warm Average Spender	421	88	2	123	10	3	647
Warm Big Spender	178	11	40	15	67	17	328
Active Average Spender	0	427	4	314	17	16	778
Active Guest Big Spender	0	16	186	47	73	65	387
Active Recurring Big Spender	0	0	16	10	0	93	119
Total	848	574	257	535	186	195	2595

QUARTER 3

QUARTER 4							
	Cold	Warm Average Spender	Warm Big Spender	Active Average Spender	Active Guest Big Spender	Active Recurring Big Spender	Total
Cold	541	54	26	157	60	10	848
Warm Average Spender	320	77	10	252	26	16	701
Warm Big Spender	131	5	52	18	63	41	310
Active Average Spender	0	257	7	411	9	34	718
Active Guest Big Spender	0	7	156	24	59	53	299
Active Recurring Big Spender	0	2	24	10	0	160	196
Total	992	402	275	872	217	314	3072



Purchase Probability Score

Logit Model



The Logit Model

Active in next quarter \sim Recency + Frequency + Monetary Value + Average Purchase Amount + Previous Segment

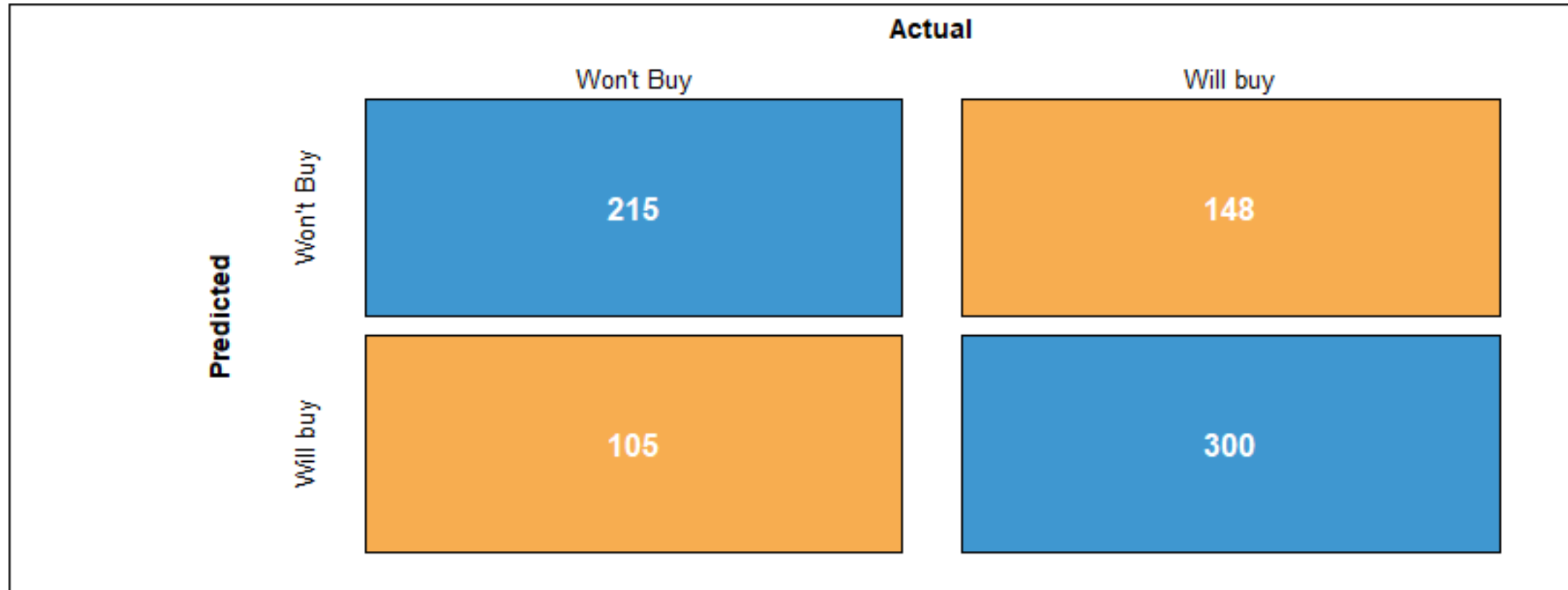
Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.2388746	0.1853431	1.289	0.19746	
recency	-0.0049678	0.0008647	-5.745	9.19e-09	***
frequency	0.0988137	0.0389480	2.537	0.01118	*
avg_amount	-0.0002848	0.0003228	-0.882	0.37761	
revenue	0.0004156	0.0001486	2.796	0.00518	**
prev_segmentactive guest big	0.1701740	0.1841677	0.924	0.35548	
prev_segmentactive recurring big	-0.2662221	0.5486469	-0.485	0.62751	
prev_segmentcold	0.2504969	0.1976134	1.268	0.20494	
prev_segmentNone	-0.5284245	0.1677057	-3.151	0.00163	**
prev_segmentwarm average	-0.0554109	0.1473408	-0.376	0.70686	
prev_segmentwarm big	-0.4427841	0.1891025	-2.342	0.01921	*

- Trained on data till Quarter 3
- Tested using observations from Quarter 4
- Standard 75-25 Train Test Split

How'd we do?

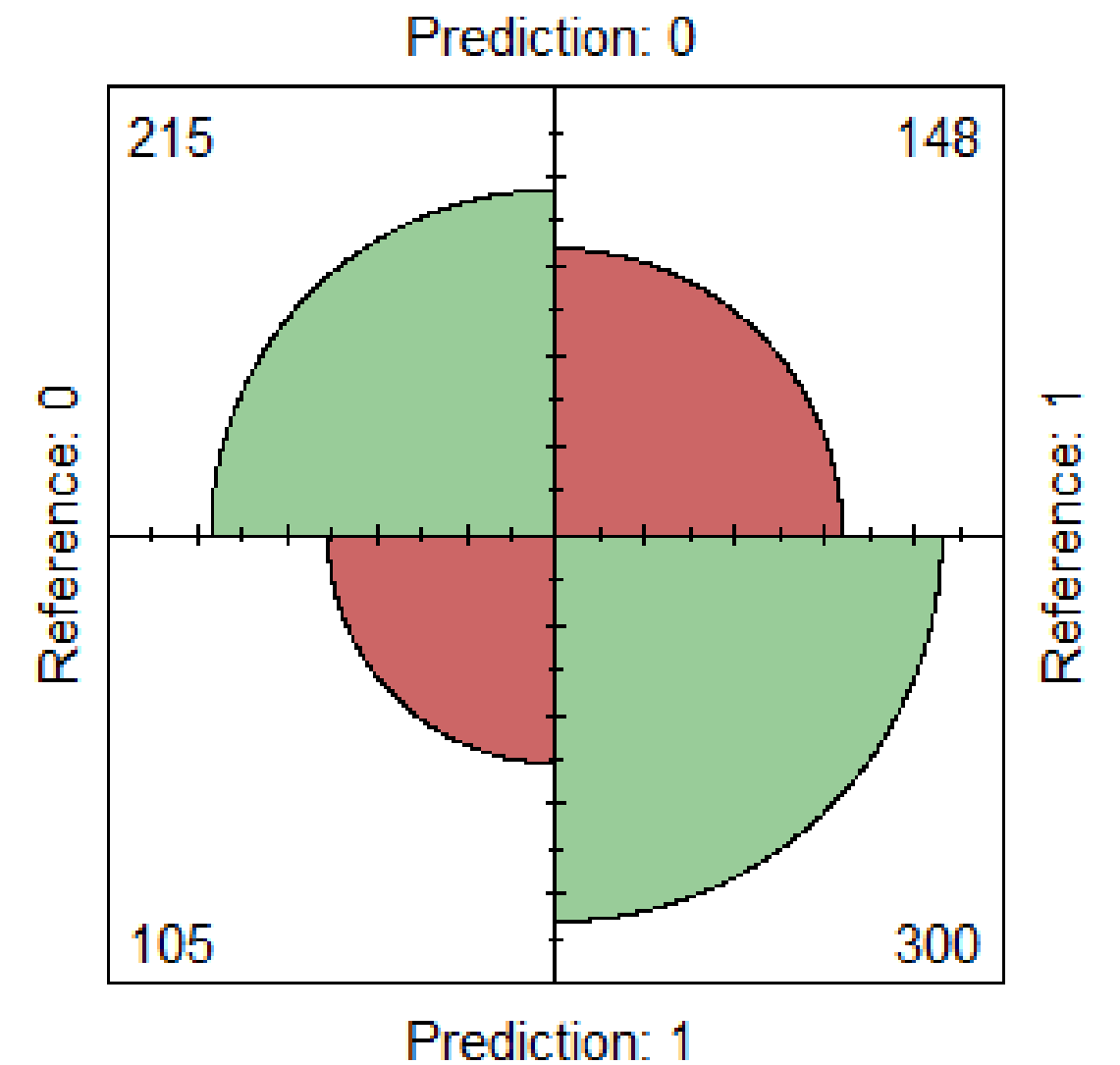
CONFUSION MATRIX



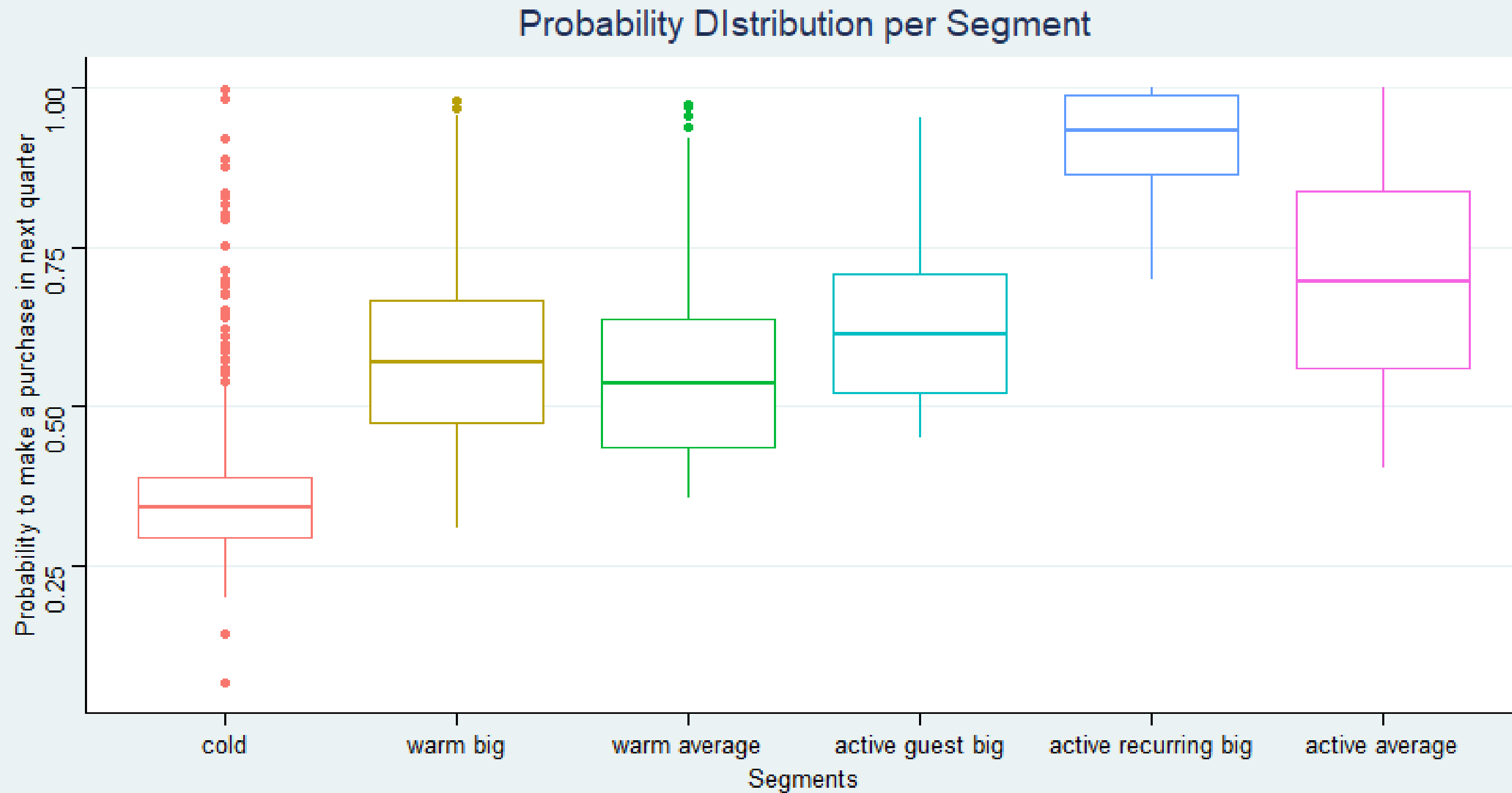
DETAILS

Sensitivity 0.672	Specificity 0.67	Precision 0.592	Recall 0.672	F1 0.63
Accuracy 0.671		Kappa 0.335		

Confusion Matrix



Probabilities for Q4





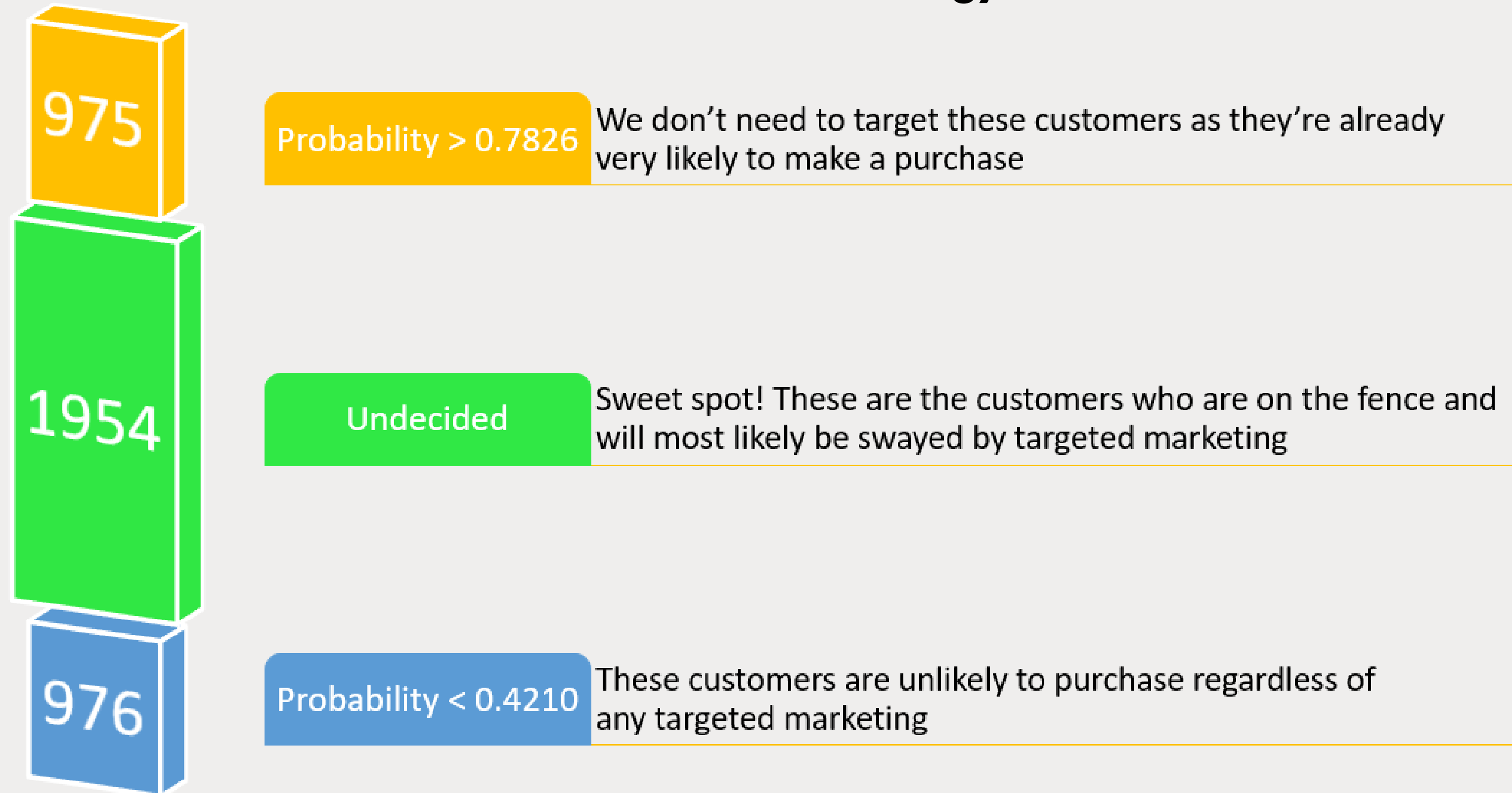
Targeted Marketing

Market Basket Analysis

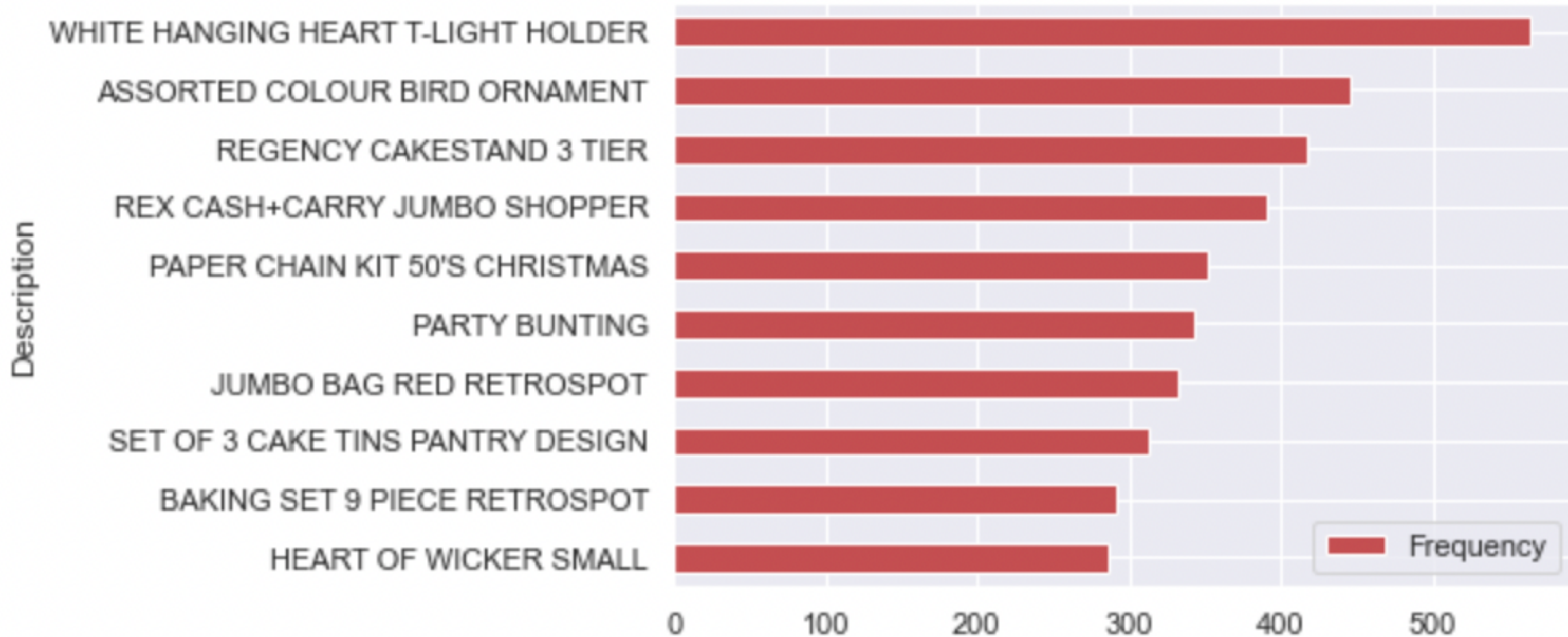


Who do we target?

The 'Middle-50' Strategy



Top 10 Items Sold



Market Basket Analysis

Rule: $X \Rightarrow Y$

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

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

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



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

- 68 Rules
- Support 0.017
- Min Threshold 1

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY...	(GREEN REGENCY TEACUP AND SAUCER)	0.021578	0.033909	0.019729	0.914286	26.963117	0.018997	11.271065
1	(PINK REGENCY TEACUP AND SAUCER)	(GREEN REGENCY TEACUP AND SAUCER)	0.027744	0.033909	0.024455	0.881481	25.995690	0.023515	8.151395
2	(PINK REGENCY TEACUP AND SAUCER, GREEN REGENCY...	(ROSES REGENCY TEACUP AND SAUCER)	0.024455	0.037402	0.019729	0.806723	21.568751	0.018814	4.980396
3	(PINK REGENCY TEACUP AND SAUCER)	(ROSES REGENCY TEACUP AND SAUCER)	0.027744	0.037402	0.021578	0.777778	20.794872	0.020541	4.331689
4	(ROSES REGENCY TEACUP AND SAUCER, GREEN REGENC...	(PINK REGENCY TEACUP AND SAUCER)	0.025483	0.027744	0.019729	0.774194	27.905376	0.019022	4.305707
...
63	(JUMBO BAG RED RETROSPOT)	(JUMBO BAG BAROQUE BLACK WHITE)	0.067201	0.030415	0.017674	0.262997	8.646913	0.015630	1.315578
64	(REGENCY CAKESTAND 3 TIER)	(ROSES REGENCY TEACUP AND SAUCER)	0.084464	0.037402	0.020962	0.248175	6.635277	0.017803	1.280348
65	(REGENCY CAKESTAND 3 TIER)	(GREEN REGENCY TEACUP AND SAUCER)	0.084464	0.033909	0.018290	0.216545	6.386109	0.015426	1.233116
66	(WHITE HANGING HEART T-LIGHT HOLDER)	(RED HANGING HEART T-LIGHT HOLDER)	0.111591	0.033703	0.023633	0.211786	6.283857	0.019872	1.225933
67	(ASSORTED COLOUR BIRD ORNAMENT)	(PAINTED METAL PEARS ASSORTED)	0.087752	0.026510	0.018290	0.208431	7.862208	0.015964	1.229823

Recommendation

- Continue updating and identifying new segments (don't just stick to the ones already identified)
- The 'Middle-50' strategy is just a suggestion. Possible to implement more granular strategies (for eg. target specific segments of interest within the middle 50 or instead, shift focus to to minimize unfavourable customer transition)
- With the focus groups identified, apply a targeted marketing strategy such as Market Basket Analysis as follows -
 - Display each consequent product as a recommendation on the antecedent product page.
 - Bundle the associated products and implement a bundle pricing technique or a sales promotion around both items.

The End

Thank you

Questions?

