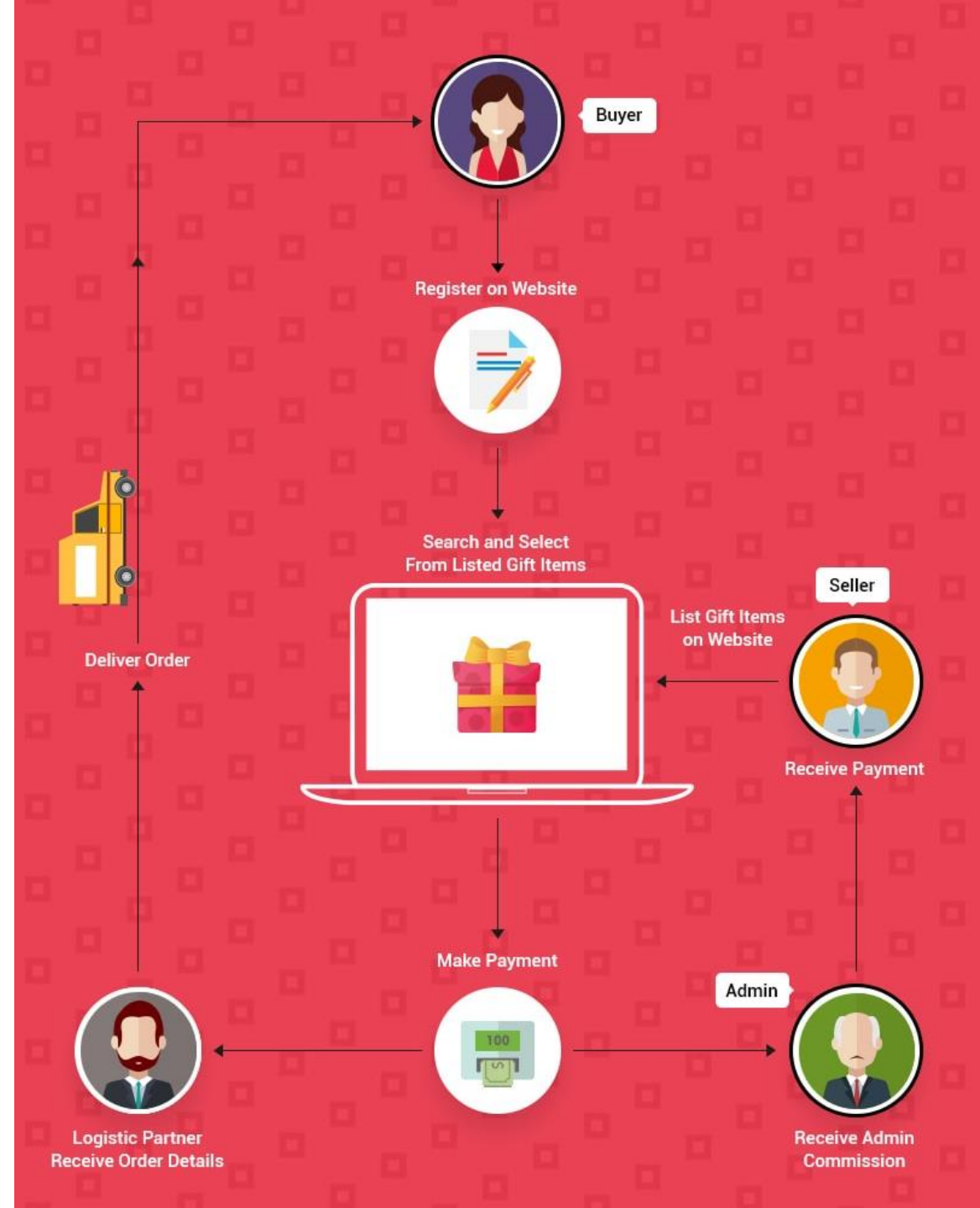


ONLINE RETAIL DATASET ANALYSIS

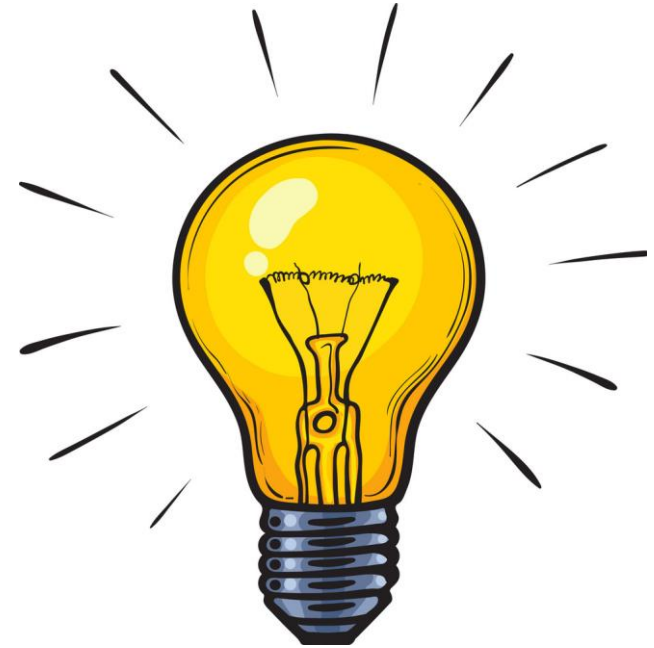
DATASET

- Transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail
- Mainly sells unique all-occasion gifts
- Customers are wholesalers



PROBLEM STATEMENT

1. How does the customer usage vary across a defined timeline (day/month)?
2. How does customer preference for products vary based on time and country?
3. How does revenue vary based on location and time?
4. Does pareto principle hold for the contribution of products to revenue?
5. How can customers be segmented based on their usage?
6. How much can we afford to spend on marketing to gain a customer?



OVERALL METRICS

OF SIGNED UP CUSTOMERS

4,219

AOV PER CUSTOMER/ DAY

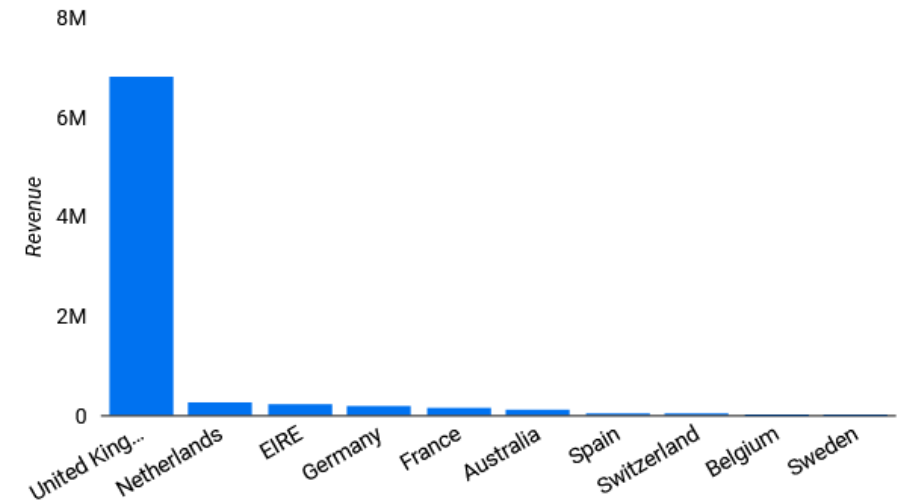
£26.92

Revenue

£8.91M

- UK contributes to almost 82% of revenue
- Focussed analysis on UK sales reveals insights with greater statistical confidence
- Region-wise analysis can inform market expansion strategies
 - i. **UK**
 - ii. **Other European Counties**
 - iii. North America → not enough customers
 - iv. Middle East → not enough customers

REVENUE CONTRIBUTION BY COUNTRY

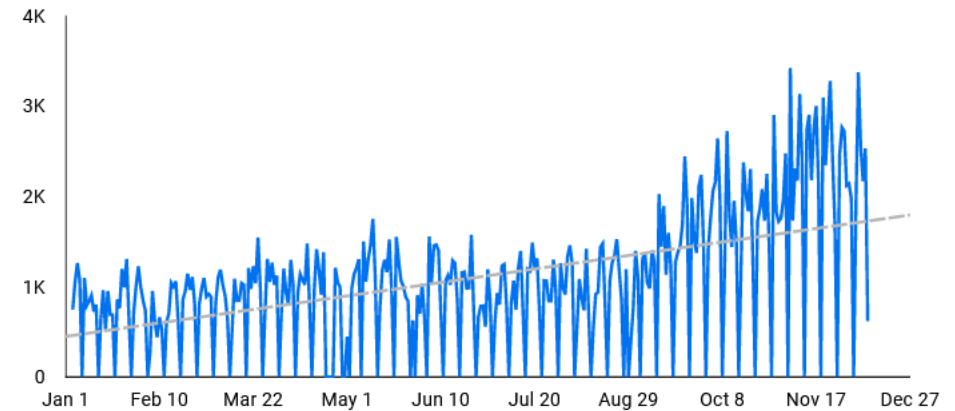


DAU/MAU

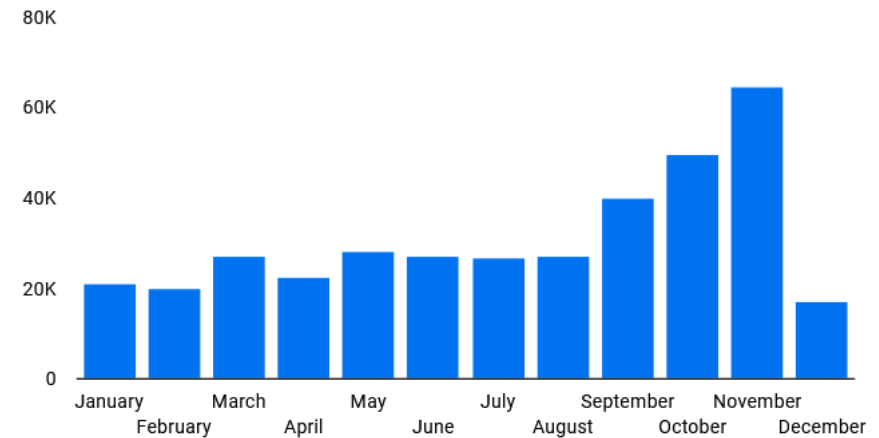
- Active user here refers to „purchasing user“
- Overall trends looks promising
- There are no users every Saturday
- Skewed by few months → Sept, Oct and Nov
- Trends are not different for UK and other European countries as expected

Q4 IS DRIVER OF REVENUE

NO OF DAILY ACTIVE USERS



NO OF MONTHLY ACTIVE USERS

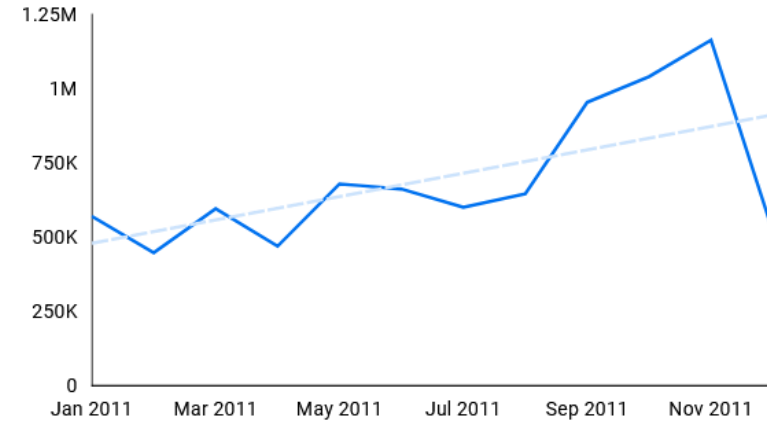


Revenue Trends

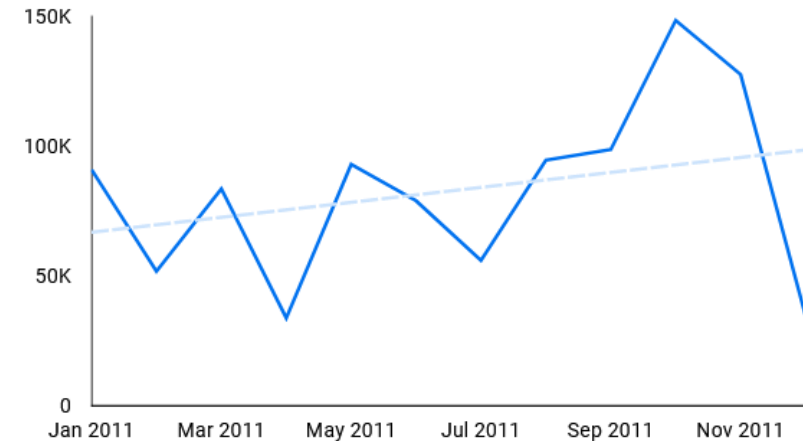
- Revenue growth is not stable
- The spike in user activity also translates to revenue growth
- Monthwise revenue tends to be cyclic
- The effects are drastic especially for other European markets

REVENUE INCREASING BUT NOT STABLE

REVENUE TREND FOR UK

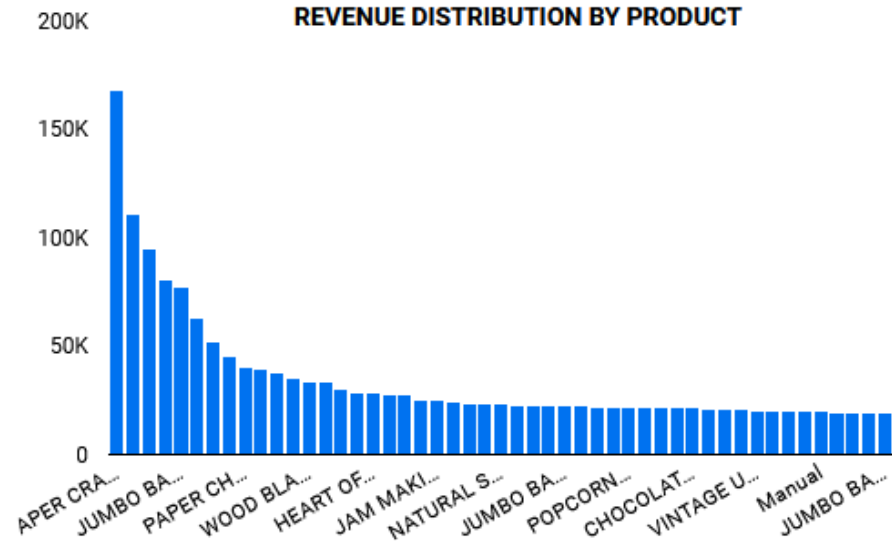


REVENUE TREND FOR OTHER EUROPEAN COUNTRIES



Product Analysis

- Total Number of Products : 3844
- Revenue distribution from the product follows Pareto principle
- In UK, the top 3 selling products are:
 1. Paper Craft Little Birde
 2. Regency Cake Stand
 3. White hanging heart t-shirt holder
- Popularity of products are highly dependent on the season (quarter)



*Most popular only
in Q4(Oct-Dec)*

AVG UNITS: 80995



*Most popular only
in Q1(Jan-Mar)*

AVG UNITS: 7



*Most popular only
in Q2(Apr-Jun)*

AVG UNITS :393



*Most popular only
in Q3(Jul-Sep)*

AVG UNITS: 7

**SEASON AFFECTS PRODUCT SALES
STRONGLY**

Cohort Analysis

- Cohort analysis provides info on monthwise customer retention
- The retention post the first month of purchase tends to stabilize
- Average Retention per month = **25%**
- However for any given month, there is no significant improvement in the customers retained
- Number of unique customer sign-ups is also stagnating for every cohort

Almost stable →

subscription_start	0	1	2	3	4	5	6	7	8	9	10	11	12
01.12.2010	100%	37%	32%	38%	36%	40%	36%	35%	35%	40%	37%	50%	27%
01.01.2011	100%	22%	27%	23%	32%	29%	25%	24%	30%	33%	36%	12%	
01.02.2011	100%	19%	19%	28%	27%	25%	25%	28%	25%	31%	7%		
01.03.2011	100%	15%	25%	20%	22%	17%	27%	23%	28%	9%			
01.04.2011	100%	21%	20%	21%	20%	23%	22%	26%	7%				
01.05.2011	100%	19%	17%	17%	21%	23%	26%	10%					
01.06.2011	100%	17%	16%	26%	23%	33%	10%						
01.07.2011	100%	18%	21%	22%	27%	11%							
01.08.2011	100%	21%	25%	24%	12%								
01.09.2011	100%	23%	30%	11%									
01.10.2011	100%	24%	11%										
01.11.2011	100%	11%											
01.12.2011	100%												

↓ *Not stable*

RETENTION CAN BE IMPROVED

CLV Analysis

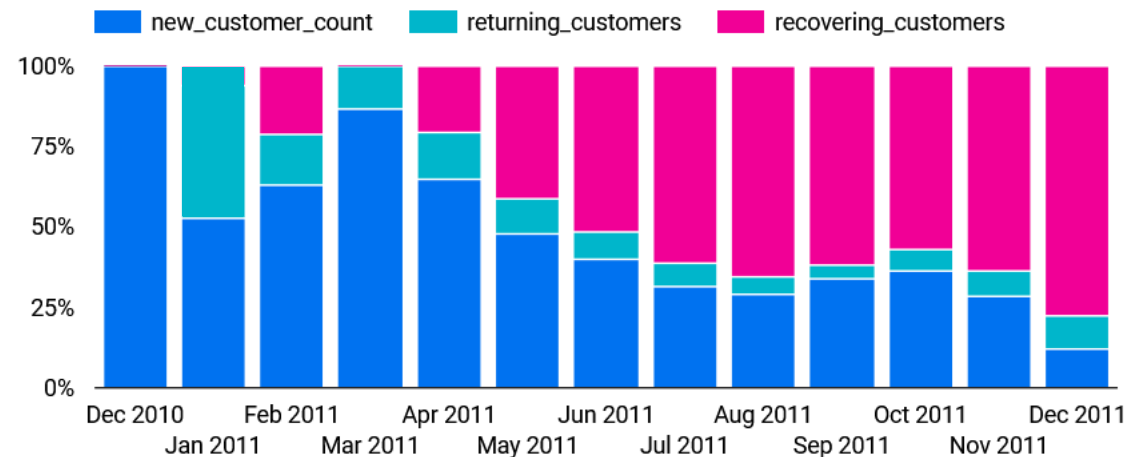
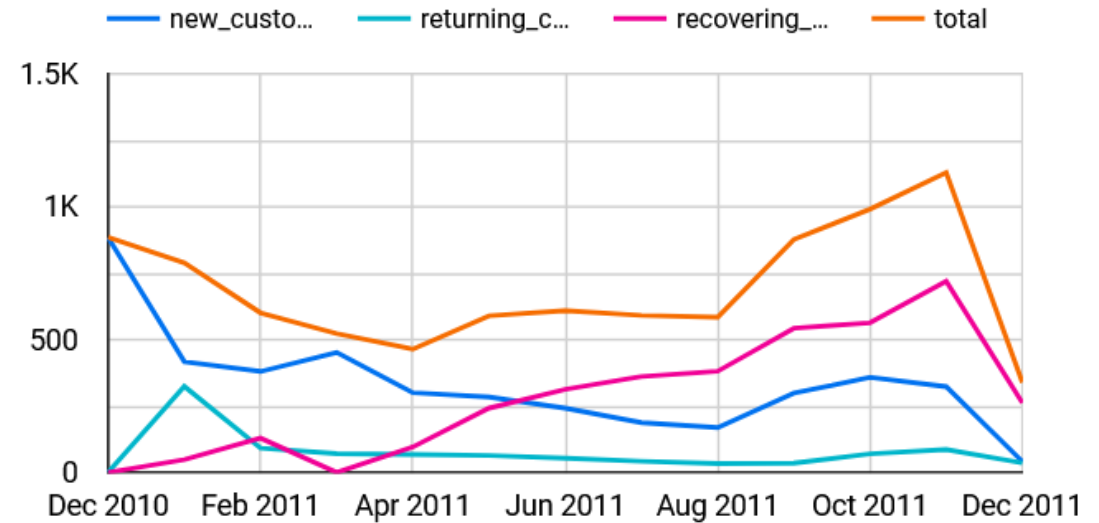
- Cohort analysis can be used to calculate CLV
- CLV calculation is done based on simple linear interpolation
- Calculated CLV = £1220
- Hence the CAC is to be set lower than this value for profitability

subscription_start	cohort_size	0	1	2	3	4	5	6	7	8	9	10	11	12
Dec 10	885	119.0	57.4	48.6	63.0	42.5	70.0	65.3	64.6	68.9	98.3	94.8	106.8	38.6
Jan 11	417	179.2	33.7	38.7	43.8	49.6	51.7	42.9	44.4	44.0	68.3	75.7	16.2	
Feb 11	380	125.0	23.0	32.5	38.1	31.7	27.1	39.3	49.4	43.8	51.2	8.4		
Mar 11	452	155.0	23.3	45.7	33.2	40.0	31.0	50.3	55.0	55.2	10.0			
Apr 11	300	156.4	37.8	32.2	31.2	33.7	38.7	36.8	43.9	8.1				
May 11	284	187.2	28.2	30.5	28.9	42.0	49.7	50.2	270.4					
Jun 11	242	248.6	27.1	25.9	56.7	49.0	78.4	15.1						
Jul 11	188	197.4	31.4	41.4	46.8	52.0	16.2							
Aug 11	169	258.9	68.1	115.3	144.6	49.6								
Sep 11	299	314.6	58.5	75.1	25.0									
Oct 11	358	360.3	82.5	26.0										
Nov 11	323	380.5	42.6											
Dec 11	41	660.0												
	333.													
	7	257.1	42.8	46.5	51.1	43.4	45.3	42.8	87.9	44.0	56.9	59.6	61.5	

**CLV CORRESPONDS TO MARKET STANDARD
FOR B2B BUSINESS**

Customer Segmentation

- New customers – customer who make their first purchase
- Returning – customers who bought last month and again buys in current month
- Recovered – customers who made their previous purchase atleast 2 months before and buys again
- % of recovered customers is increasing → positive sign
- Absolute number of new customers is decreasing → concern

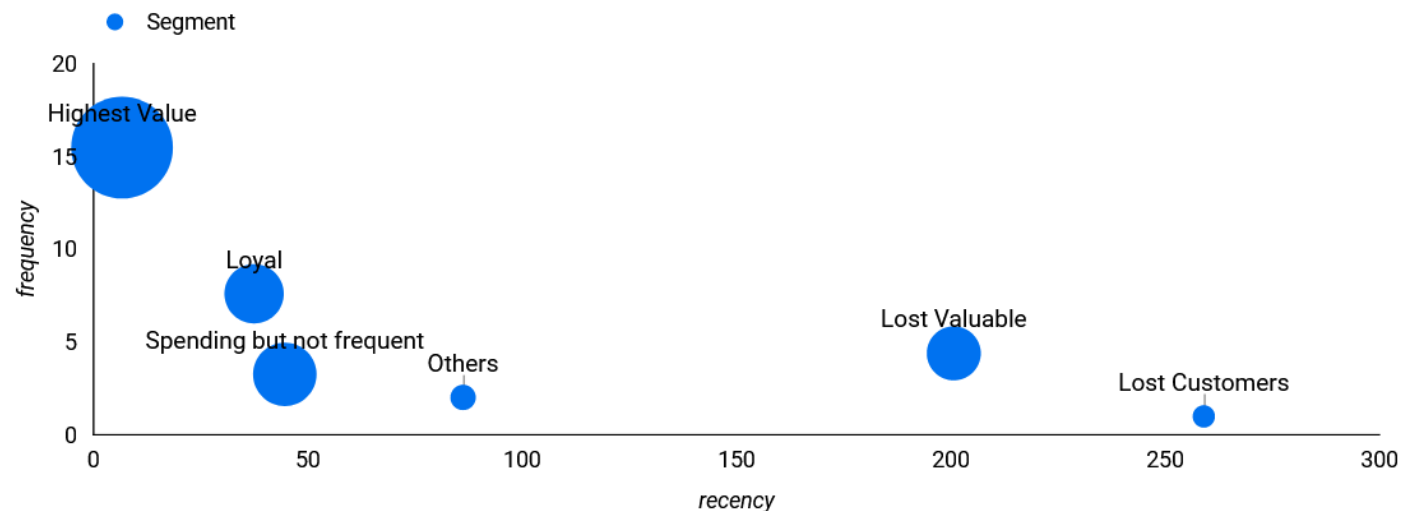


Customer Segmentation

- Highest Value customers – 11%
- Loyal customers – 13.8%

These customers need more focus with respect to pricing plans and product variety

- Spending but not frequent customers are customers that need to be brought to the platform more often
- Loyal customers tend to bring high monetary value but the recency is almost **6 x** highest value customers



	Segment	frequency ▾	monetary	recency	Record Count
1.	Highest Value	15.48	9,184.41	6.67	483
2.	Loyal	7.61	3,130.11	37.47	572
3.	Lost Valuable	4.4	2,589.92	200.7	93
4.	Spending but not frequent	3.27	3,583.55	44.63	226
5.	Others	2.02	574.14	86.22	2,441
6.	Lost Customers	1	444.93	259.03	523

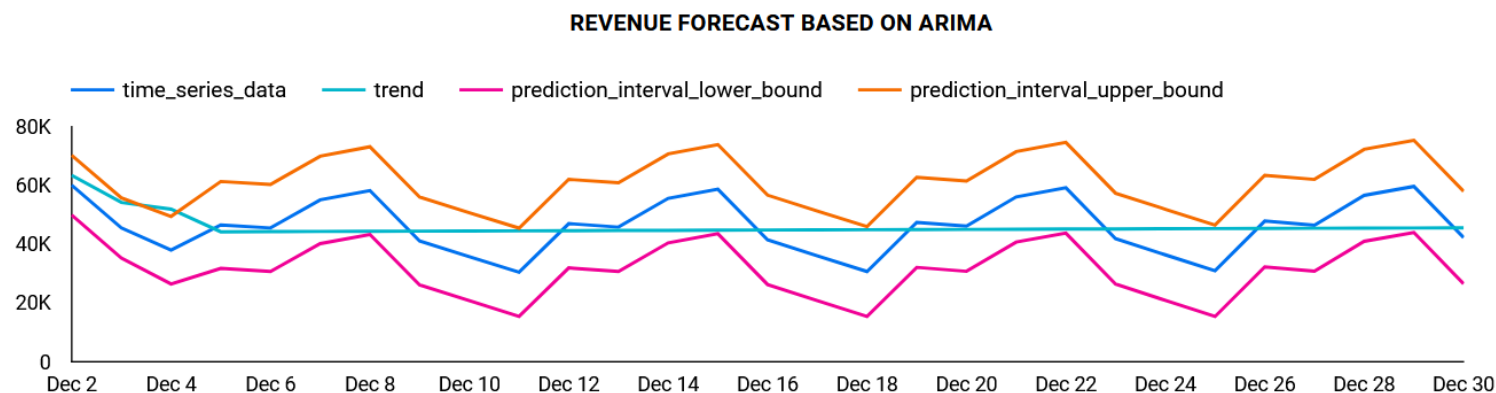
**CUSTOMER SEGMENTATION EMPHASIZES
THE NECESSITY OF MORE FREQUENT BUYERS**

Revenue Forecast

- Forecasting revenue for next 30 days is crucial to assess the business performance and discover underlying patterns
- Forecasting was done using ARIMA model trained on data till December 2011
- Forecasting accuracy of 92% achieved till 9th December 2011 for which we have data.
- With time, better models can be utilized for more accurate forecast over longer periods of time

NEXT 30 DAY REVENUE FORECAST

£1.34M



Next Steps and Recommended Strategies

From **Customer Segmentation** and **cohort analysis**, we understand that more customers do **not** have a stable purchase pattern

- A subscription-based plan will be beneficial to improve customer retention and increase the number of returning customers.

(The recency of the most loyal customers are 6X highest value customers)

- Greater discounts for highest value customers to retain them through reward point programs
- Industry Standard CAC is £580 for B2B business which provides leeway for better marketing campaigns and discounts

From **Product Analytics**, the contribution of products to revenue follows pareto distribution

- Seasonal Discounts for products will prove beneficial
- The minimum order quantity of products per purchase can be increased by providing discount sale

(since the average number of units sold is not high on a yearly basis)