

CLVT Analysis

Onur AKCAKAYA

CLTV-Final Project

- Business Problem
- Data Story
- Variable Description
- Aim
- Method
- Results
- Recommendations



1-Business Problem



An UK based e-commerce company wants evaluate their UK based customers according to their future purchases. Following tasks are asked:

- 6 months of purchases and customer evaluation
- Comparison of 1-month and 12-months prediction
- Customer segmentation of 3 groups for 6-months prediction (C,B,A)
- Recommend 6-months action for each segment

2-Data Story

- The data set “Online Retail II” contains transaction records of an UK based e-commerce store between 12.01.2019-12.09.2011
- Items listed are souvenirs also can be considered as promotional items
- Majority of the customers are wholesale traders

3-Variables

- **InvoiceNo:** Invoice number. Nominal. A 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.\n
- **StockCode:** Product (item) code. Nominal. A 5-digit integral number uniquely assigned to each product.\n
- **Description:** Item name and description. Nominal.\n
- **Quantity:** The quantities of each item per transaction. Numeric.\n
- **InvoiceDate:** Invoice date and time. Numeric.\n
- **UnitPrice:** Unit price. Numeric. Product price per unit (£).\n
- **CustomerID:** Customer number. Nominal. A 5-digit number uniquely assigned to each customer.\n
- **Country:** Country name. Nominal. The name of the country where a customer resides.

4-AIM

- Aim of the project is to define customer value in a specific time according to their future purchases



5-METHOD

- Dataset obtained from: <https://archive.ics.uci.edu/ml/machine-learning-databases/00502/>
- **Missing values** and “**Cancelled**” transactions are dropped from the data set
- **TotalPrice** column created for each invoice by using “**Price**” and “**Quantity**” Columns
- **recenc_cltv_p**, **T**, **frequency** & **monetary_avg** values are calculated and columns are created
- **BG/NBD model** deployed and **expected_number_of_purchases** are calculated
- **Gamma-Gamma model** deployed and **expected_average_profit** are calculated then CLTV are calculated.
- Interpretations and recommendations are based on CLTV prediction model.

7-Result

	Customer ID	recency_cltv_p	T	frequency	monetary_avg	recency_weekly_p	T_weekly	expected_number_of_purchases_6m	expected_average_profit_6m	clv_6
2486	18102	366	368	60	3584.88775	52.28571	52.57143	22.73601	3595.19256	85651.01047
589	14096	97	102	17	3159.07706	13.85714	14.57143	16.66160	3191.38673	55650.64677
2184	17450	359	368	46	2629.52989	51.28571	52.57143	17.54836	2639.41934	48533.31011
2213	17511	370	374	31	2921.95194	52.85714	53.42857	11.95148	2938.27449	36797.00673
1804	16684	353	359	28	2120.04696	50.42857	51.28571	11.22187	2133.20359	25083.02541
406	13694	369	374	50	1267.36260	52.71429	53.42857	18.80521	1271.78540	25060.70871
587	14088	312	323	13	3859.60154	44.57143	46.14286	6.10321	3911.31885	25010.05913
1173	15311	373	374	91	667.59681	53.28571	53.42857	33.65832	668.89446	23591.38948
133	13089	366	370	97	605.18660	52.28571	52.85714	36.08923	606.29366	22927.69296
1057	15061	368	373	48	1108.30781	52.57143	53.28571	18.12245	1112.34712	21123.08206
1485	16000	0	3	3	2055.78667	0.00000	0.42857	9.26034	2181.32360	21110.80764
692	14298	352	361	44	1159.85545	50.28571	51.57143	17.07954	1164.46444	20839.63143
1378	15769	363	372	26	1727.11615	51.85714	53.14286	10.16870	1738.69116	18526.06343
2499	18139	0	19	6	1406.39000	0.00000	2.71429	11.76767	1448.27469	17817.89132
2377	17841	371	374	124	328.82234	53.00000	53.42857	45.54853	329.30652	15717.33660
454	13798	370	373	57	650.51114	52.85714	53.28571	21.39602	652.53376	14629.74945
2150	17389	330	332	34	928.76485	47.14286	47.42857	14.34210	933.56861	14028.56200
2424	17949	370	372	45	768.52378	52.85714	53.14286	17.09377	771.53780	13819.55167
1503	16029	335	374	63	889.76040	47.85714	53.42857	14.54030	892.24077	13594.32494
139	13098	316	318	28	1031.51571	45.14286	45.42857	12.36760	1037.98629	13449.72336

[+ Code](#)[+ Markdown](#)

7-Result

- The model predicts customer values according to their expected number of purchases and expected average profit.
- Frequent buyers with high average monetary values have increased CLV's since they are loyal and dependable customers.
- Two customers, 16000 and 18139 are noticeable in the table. Even though their frequency and recency score are low, because of the high monetary value their CLVs are high. They are potential “Champions” but they need to be monitored and encouraged to make more frequent purchases
- Longtime frequent buyers should be rewarded since they are the most valuable type of customer according to model

7-Result

Out[51]:

	Customer ID	recency_cltv_p	T	frequency	monetary_avg	recency_weekly_p	T_weekly	expected_average_profit	expected_number_of_purchases_1m	clv_1
2486	18102	366	368	60	3584.88775	52.28571	52.57143	3595.19256	3.85079	14884.97500
589	14096	97	102	17	3159.07706	13.85714	14.57143	3191.38673	2.87354	9855.87989
2184	17450	359	368	46	2629.52989	51.28571	52.57143	2639.41934	2.97227	8434.76472
2213	17511	370	374	31	2921.95194	52.85714	53.42857	2938.27449	2.02407	6394.32432
1804	16684	353	359	28	2120.04696	50.42857	51.28571	2133.20359	1.90146	4361.05396
587	14088	312	323	13	3859.60154	44.57143	46.14286	3911.31885	1.03575	4355.48526
406	13694	369	374	50	1267.36260	52.71429	53.42857	1271.78540	3.18452	4354.46853
1173	15311	373	374	91	667.59681	53.28571	53.42857	668.89446	5.69939	4098.86945
133	13089	366	370	97	605.18660	52.28571	52.85714	606.29366	6.11174	3984.05150
1485	16000	0	3	3	2055.78667	0.00000	0.42857	2181.32360	1.64094	3843.97950

+ Code

+ Markdown

Out[52]:

	Customer ID	recency_cltv_p	T	frequency	monetary_avg	recency_weekly_p	T_weekly	expected_number_of_purchases_12m	clv_12
2486	18102	366	368	60	3584.88775	52.28571	52.57143	44.74416	163591.12676
589	14096	97	102	17	3159.07706	13.85714	14.57143	32.36049	104900.44292
2184	17450	359	368	46	2629.52989	51.28571	52.57143	34.53383	92694.27482
2213	17511	370	374	31	2921.95194	52.85714	53.42857	23.52180	70285.65230
1804	16684	353	359	28	2120.04696	50.42857	51.28571	22.07650	47890.36421
406	13694	369	374	50	1267.36260	52.71429	53.42857	37.01346	47871.90000
587	14088	312	323	13	3859.60154	44.57143	46.14286	11.99122	47688.86365
1173	15311	373	374	91	667.59681	53.28571	53.42857	66.25203	45067.80939
133	13089	366	370	97	605.18660	52.28571	52.85714	71.02985	43795.47286
1057	15061	368	373	48	1108.30781	52.57143	53.28571	35.66849	40348.81669

7-Result

When 1-month and 12-months CLTV compared:

- Top 10 customers are almost the same except two lines
- Customers 14088 & 13694 interchanged their orders. When making predictions in longer time periods, frequent buyers and higher tenure customers are more stable even their monetary average is low
- In the first table, 10th row customer 16000 is a rising star in short term (1month) with a higher average monetary value, in longer term a loyal-long term customer 15061 is providing higher value.

7-Recommendations



For the next 6-months:

- Segment C: predicted median frequency is 4 and average recency as high as 112 day. Sales needs to be improved. An advance recommendation system can make customers in this segment buy more expensive items and with help of customized discount rates sales increase in quantity can be maintained.
- Segment B: Monetary value is acceptable for this group. A subscription program (regular amount of sale in a time period) with a reasonable discount can maintain better recency and frequency
- Segment A: This group consists of regular high payer. Adding promotions and occasional discounts will sustain their satisfaction and hence sales.