Working of Customer life time value

Purpose:

It is used to identify the active customer from the inactive customer. This is important so that company can spend its money on engaging with its active customer and profitable customer, thus generating more revenue.

Concept: customers relation with company=15 years

   Customers per year spend=$100

   Money spent in acquiring that customer=$50

   Customer life time value= (100*15)-50

                =$1450

The CLV models needs to be built every 3-6 month to factor in the latest customer specifications.

Ways to improve CLV: personalized greetings, sending coupon codes, giving discounts.

The project will be using :1. BG/NBD model .

                2. Gama Gama model

1) BG/NBD model: this model is used to calculate the expected number of transaction per customer. This model can also predicts if the customer is active or dead.

2) Gama Gama model: this model is used to calculate Revenue per transaction

CLV= expected no of transaction * revenue per transaction * margin

Margin is usually proposed by the domain expert based on the historical purchases. We will take standard value of 5%.

Data set: the data set has been used from Kaggle.

Report summery

### 1) Analysis of Frequency and Recency.

This is calculated using summary_data_from_transaction_data.

Frequency: represents the number of repeat purchases the customer has made.

Recency: recency represents the age of the customer when they made their most recent purchases. This is equal to the duration between a customer's first purchase and their latest purchase.

customer_unique_id	frequency	recency	T	monetary_value
32ea3bdedab835c3aa6cb68ce66565ef	1.0	566.0	661.0	56.78
0b3dc7efaafb0cf78a4796d42fa8d74c	1.0	474.0	630.0	356.13
4962136f755981b83b03f476e66eea7d	1.0	427.0	477.0	163.53
8329519e31cb1b89bd44c3c6ae417ad6	1.0	389.0	477.0	111.30
2dc7f51f05ba1b1a92608e45e55d1fe1	1.0	352.0	362.0	67.43
...	...	...	...	...
558cb7d841f132be98d632d22e6bde5f	0.0	0.0	45.0	0.00
558d25e550c5a1300c5775a26df16158	0.0	0.0	184.0	0.00
558e83cbe6bde610b528833d16ba588e	0.0	0.0	166.0	0.00
558e8c5c045f205b5b9ca29ec2d281b3	0.0	0.0	171.0	0.00
ffffd2657e2aad2907e67c3e9daecbeb	0.0	0.0	329.0	0.00

2) Frequency graph: to see what is the frequency distribution of the purchase by the customers
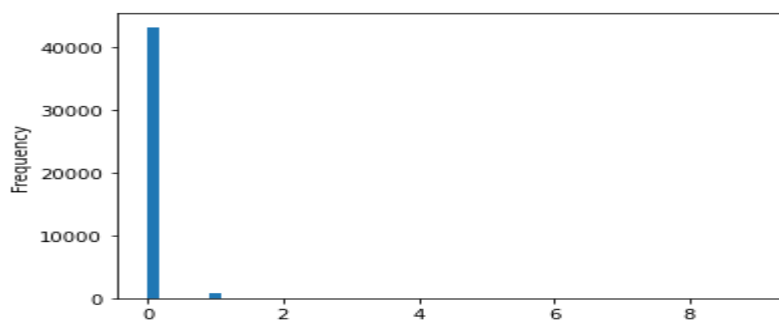
This graph tell us about the number of transactions done by the customers

We can see majority of the customers have not done repeated transactions.
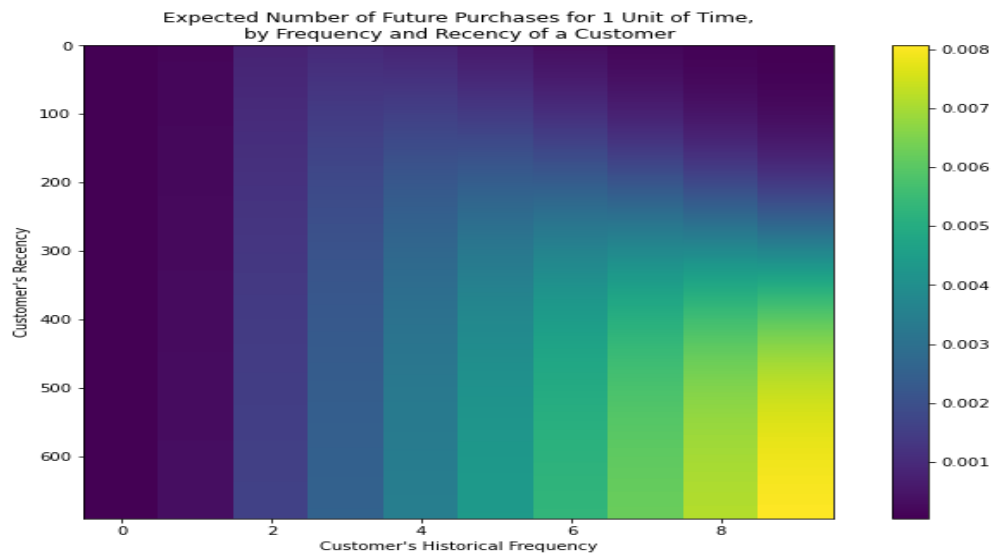
3) Fitting the bg/nbd model.

It is done using BetaGeoFitter

Bgf=BetaGeoFitter(penalizer_coef=0.0)

penalizer_coef is used to control the scaling of the model.it is used to control the parameter size. If the sample size is small then parameter will be large
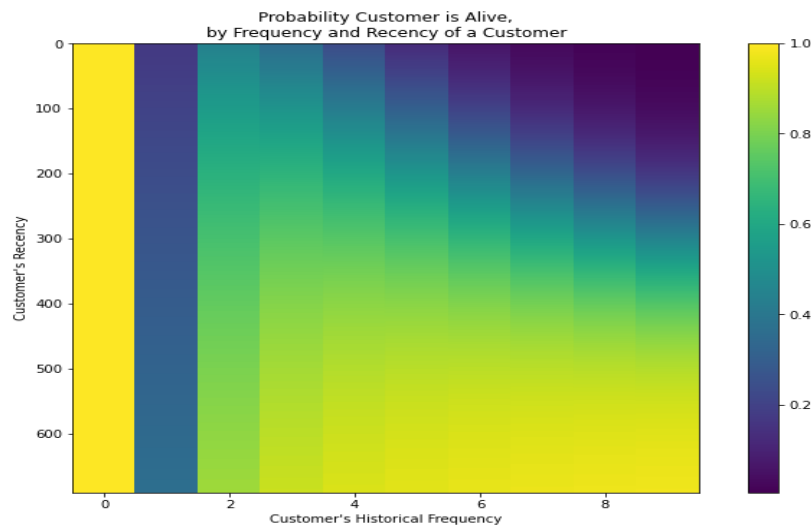
4) Frequency recency matrix.

It is done using plot_frequency_recency_matrix

5) Probability of alive matrix

It is done using plot_probability_alive_matrix

6) Predict transaction in the next 10 days.

This is done using the object: conditional_expected_number_of_purchases_up_to_time

	customer_unique_id	frequency	recency	T	monetary_value	pred_num_txn
0	8d50f5eadf50201ccdcedfb9e2ac8455	9.0	198.0	237.0	46.850000	0.13
1	1b6c7548a2a1f9037c1fd3ddfed95f33	4.0	252.0	263.0	135.820000	0.06
2	dc813062e0fc23409cd255f7f53c7074	4.0	322.0	358.0	158.987500	0.05
3	47c1a3033b8b77b3ab6e109eb4d5fdf3	3.0	142.0	176.0	252.853333	0.05
4	fc719c11b4dce594f47e1f9aa739a3a2	2.0	40.0	45.0	101.480000	0.04
5	63cfc61cee11cbe306bff5857d00bfe4	3.0	168.0	224.0	157.273333	0.04
6	b8b3c435a58aebd788a477bed8342910	3.0	155.0	240.0	299.380000	0.04
7	ec7f1811826ab04a27a92197bc40c888	3.0	311.0	323.0	85.833333	0.04
8	a7657330b1c135f3acd420326e335b2c	3.0	179.0	295.0	197.926667	0.04
9	f0e310a6839dce9de1638e0fe5ab282a	3.0	219.0	298.0	130.856667	0.04

7) Accessing model fit

This graph will tell us about how well our model fits the real data.
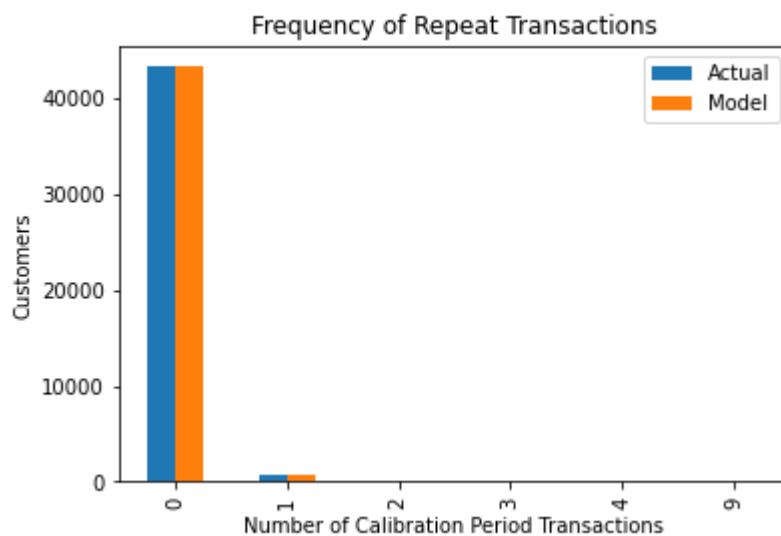
We can see our model fits well with the real data.

8) Listing customer who have made repeated purchase

```
             customer_unique_id  frequency   recency      T  \
0   013f4353d26bb05dc6652f1269458d8d       1.0      4.0   37.0
1   015557c9912277312b9073947804a7ba       1.0     77.0  360.0
2   018b5a7502c30eb5f230f1b4eb23a156       1.0     94.0  227.0
3   01c289bb06354cdc7e6549570f20ada4       1.0    295.0  300.0
4   02168ea18740a0fdaaa15f11bebba5db       1.0     21.0  133.0

   monetary_value  pred_num_txn
0          164.04          0.01
1          242.22          0.00
2           55.09          0.01
3          168.36          0.01
4          132.02          0.01
------------------------------------------
The Number of Returning Customers are:  754
```

ase

7) Fitting Gama Gama model:

It relies on important concept that there is no relation between monetary value and purchase frequency. So we need to check if the Pearson correlation between them is close to zero.

8) Calculating life time value

	customer_unique_id	CLV
0	86df00dc5fd68f4dd5d5945ca19f3ed6	697.25
1	a1044dd75b74fbc485b040575a14acf0	404.68
2	b8b3c435a58aebd788a477bed8342910	391.16
3	47c1a3033b8b77b3ab6e109eb4d5fdf3	376.26
4	3db6e5afa8e8178a29afc729855315e2	330.13
5	21bde4e982c4681021edabb899bcebfe	300.75
6	055ec572ac7f3c7bdd04a183830ebe59	290.18
7	6419a1be8feac26ec793667b71cbaeb4	270.78
8	e30b83af13d6ff0b0f427b2a67c43b39	268.58
9	dc813062e0fc23409cd255f7f53c7074	254.16