Customer Life Time Value(CLV) Prediction

- Dataset Contains a European Retail store Transactional data

Descriptive

- Cohort based CLV

Probabilistic Models

- BG/NBD Model
- GammaGamma Model

Customer Segmentation

- KMeans Clustering

Devarapu Lokesh

* For More details Visit GitHub Repo

Descriptive Method

(Describes CLV by Historical Data (by Cohort is Effective))

CLV is the Total Contribution of a Customer to a company for the total relationship with the company.

CLV= Avg Frequency * Avg Monetary * Customer Lifetime

(No of orders Annually) (Avg Purchase value) (Total no of years with company)

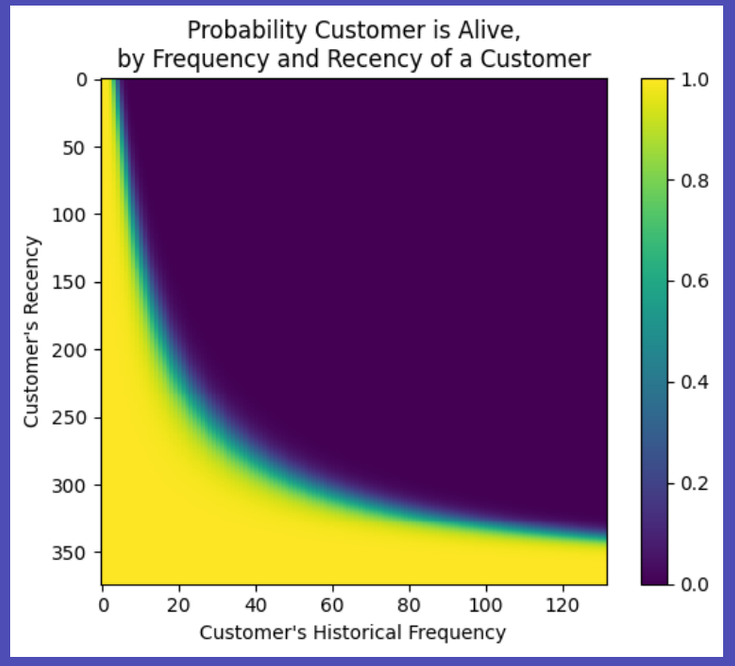
	Start_month	cohort_size	Avg_sales	Avg_Frequency	Churn_rate	Avg_lifespan	CLV_Churn	CLV_Lifetime
0	2010-12-01	885	36.422175	9.398870	0.125424	267.748023	272.936620	68.465457
1	2011-01-01	417	229.024754	5.163070	0.184652	208.083933	640.377007	236.494146
2	2011-02-01	380	36.631088	4.107895	0.228947	171.002632	65.725435	30.095331
3	2011-03-01	452	28.076714	3.564159	0.278761	140.685841	35.898084	20.013976
4	2011-04-01	300	27.500523	3.080000	0.333333	113.050000	25.410484	16.940322
5	2011-05-01	284	241.196904	2.883803	0.306338	99.954225	227.057775	139.112862
6	2011-06-01	242	83.920088	2.731405	0.355372	82.946281	64.501370	45.843949
7	2011-07-01	188	29.477879	2.351064	0.393617	57.186170	17.607058	13.860875
8	2011-08-01	169	32.904727	2.118343	0.455621	42.366864	15.298561	13.940701
9	2011-09-01	299	35.094298	1.989967	0.521739	26.655518	13.385325	13.967296
10	2011-10-01	358	25.035718	1.698324	0.653631	11.645251	6.505007	8.503752
11	2011-11-01	323	25.476888	1.359133	0.736842	3.597523	4.699308	6.925296
12	2011-12-01	41	153.916796	1.048780	0.975610	0.000000	16.546056	32.284987

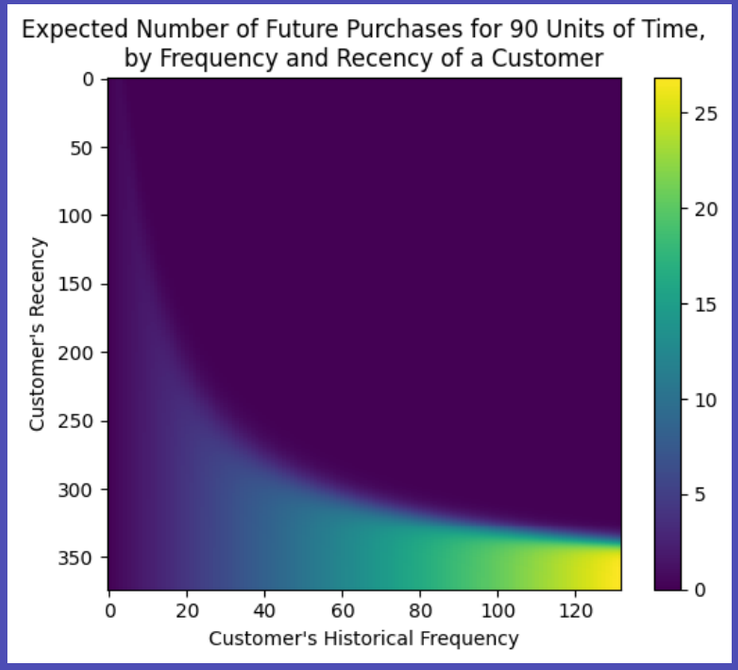
Probabilistic Method

(Predicts CLV by Probabilistic Models BG/NBD & Gamma Models)

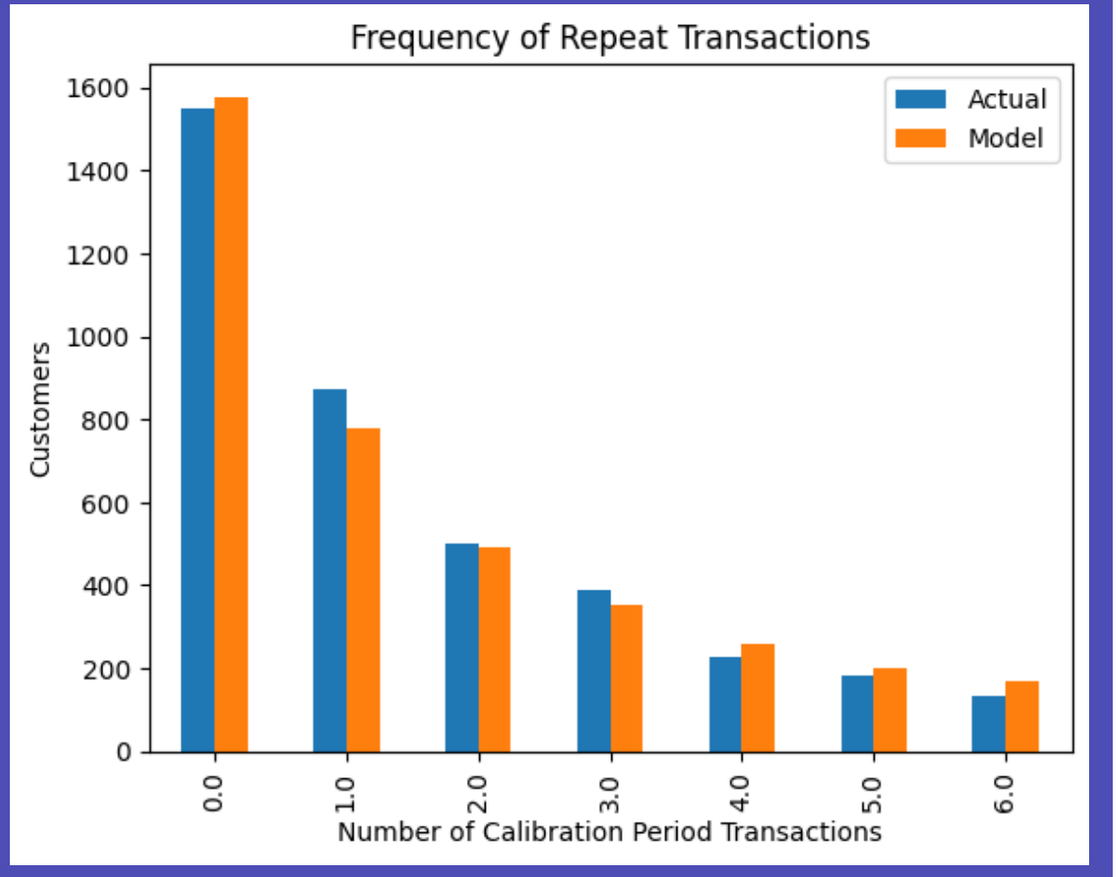
1 . BG/NBD Model

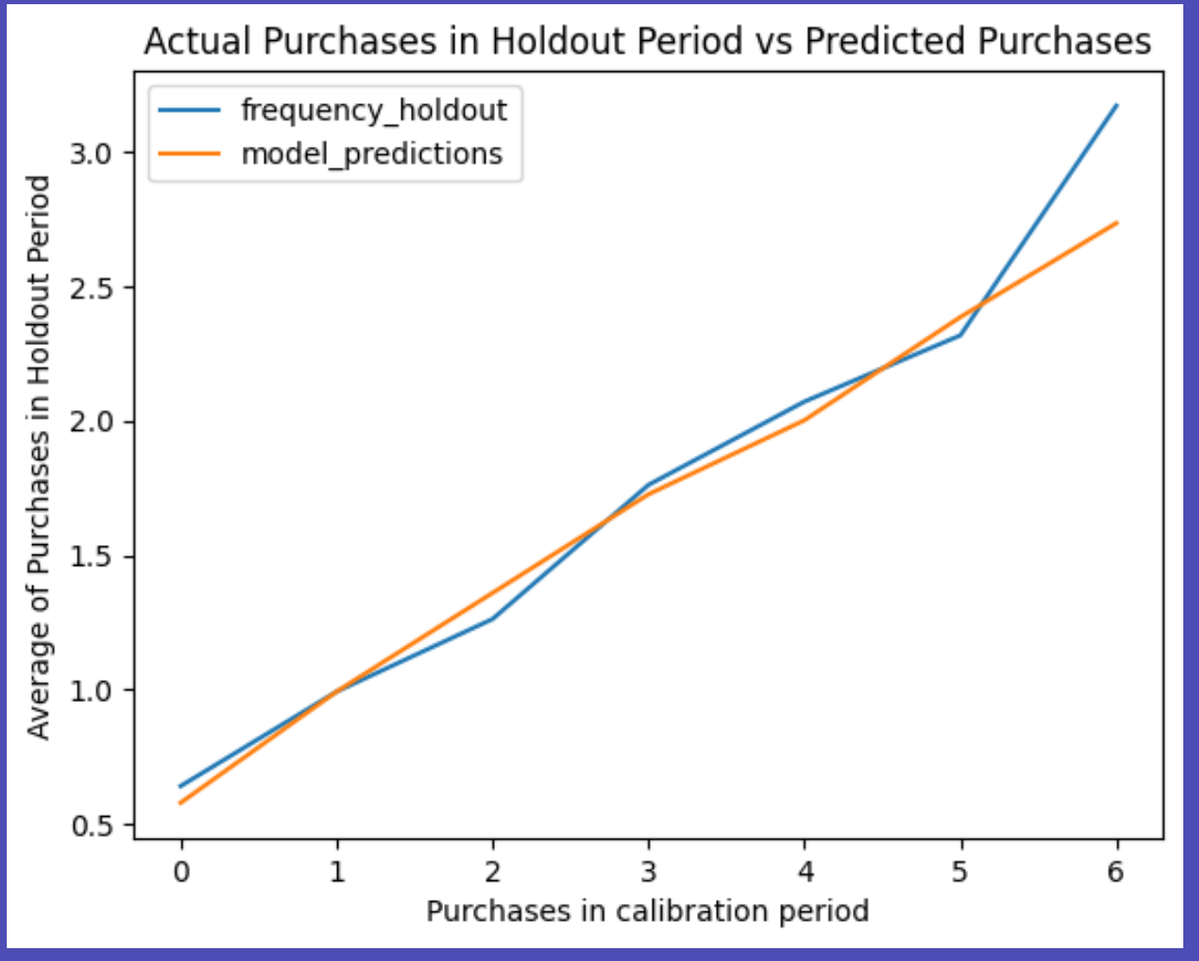
- Beta GeoMetric / Negative Binomial Distribution (BD/NBD Model)
- It Can predict Buying & Churn Behavior of Customers.
- It takes Recency, Frequency , Lifetime as input.





BG/NBD Model

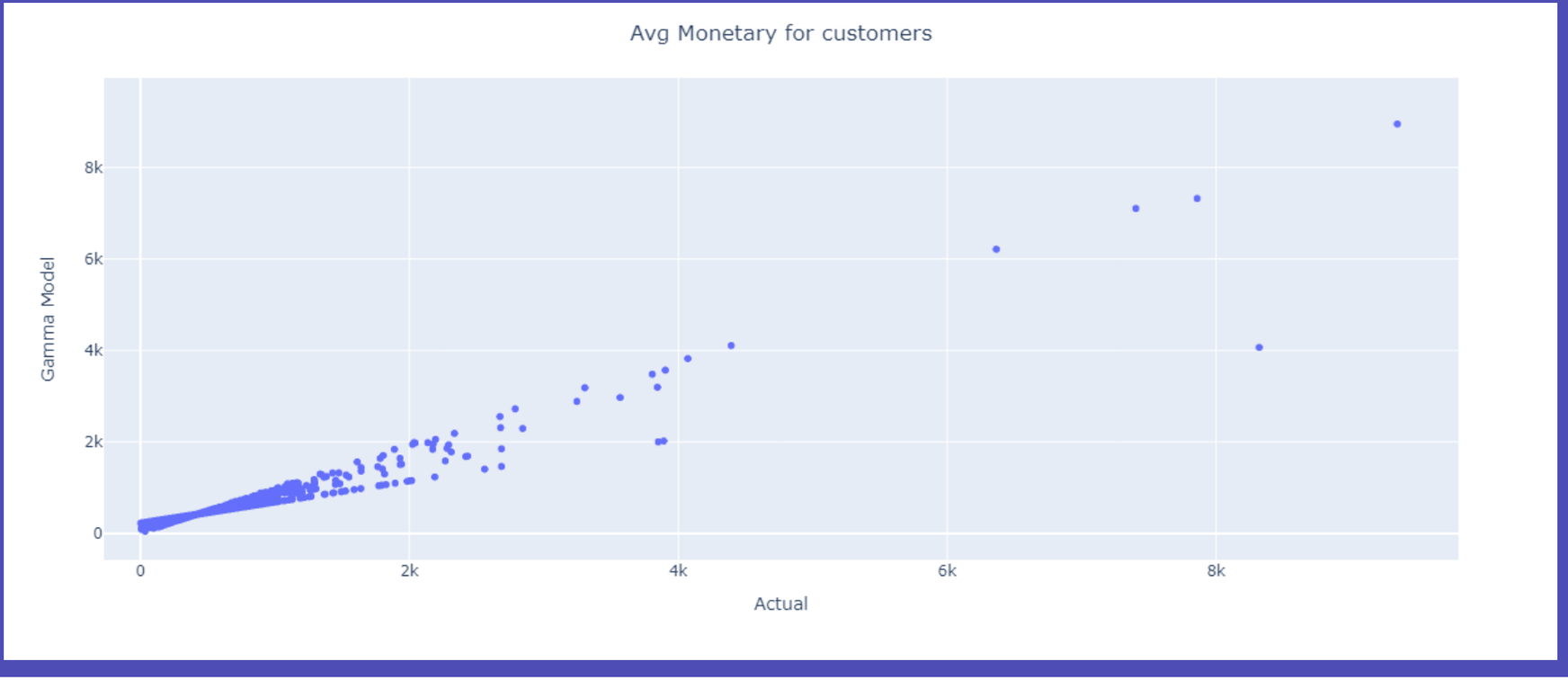

Frequency of Repeat Transactions
Actual
Model
Customers
1600
1400
1200
1000
800
600
400
200
0
0.0
1.0
2.0
3.0
4.0
5.0
6.0
Number of Calibration Period Transactions


Actual Purchases in Holdout Period vs Predicted Purchases
frequency_holdout
model_predictions
Average of Purchases in Holdout Period
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
5
6
Purchases in calibration period

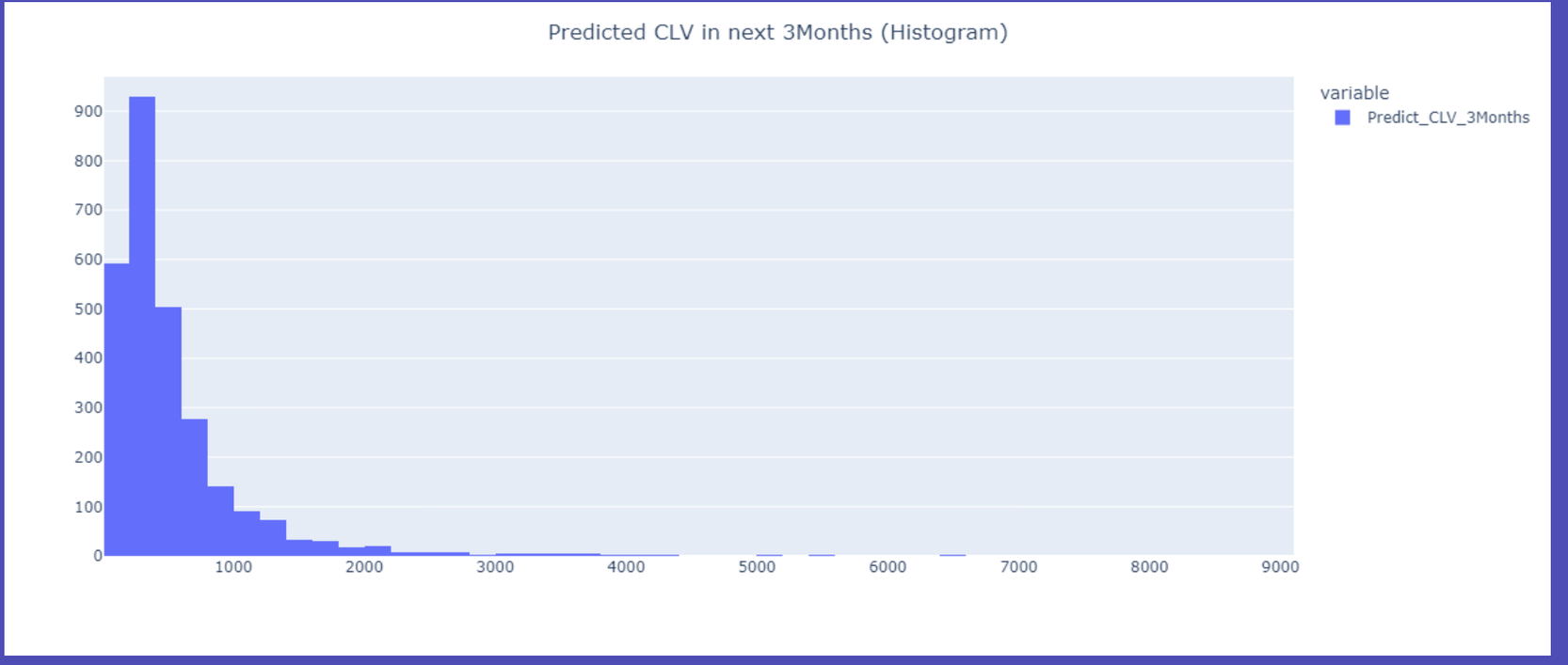
2 . Gamma-Gamma Model

- Gamma Model can Predict CLV of Customers for next N Months.
- By Combining BG/NBD and Gamma Model can answer following questions :

1. Which Customers are ALive or Churned and Which one of them will order again in next period.
2. what is the CLV of customers for next N Months.
3. The Number of Orders each Customer will place and avg Order value.



Gamma Model

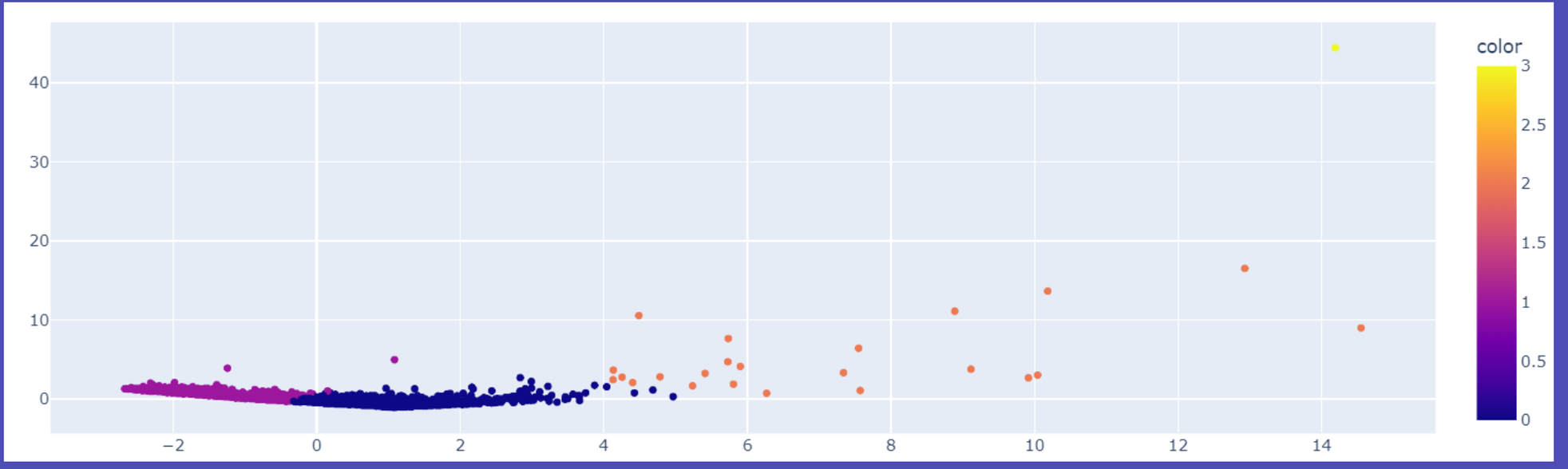


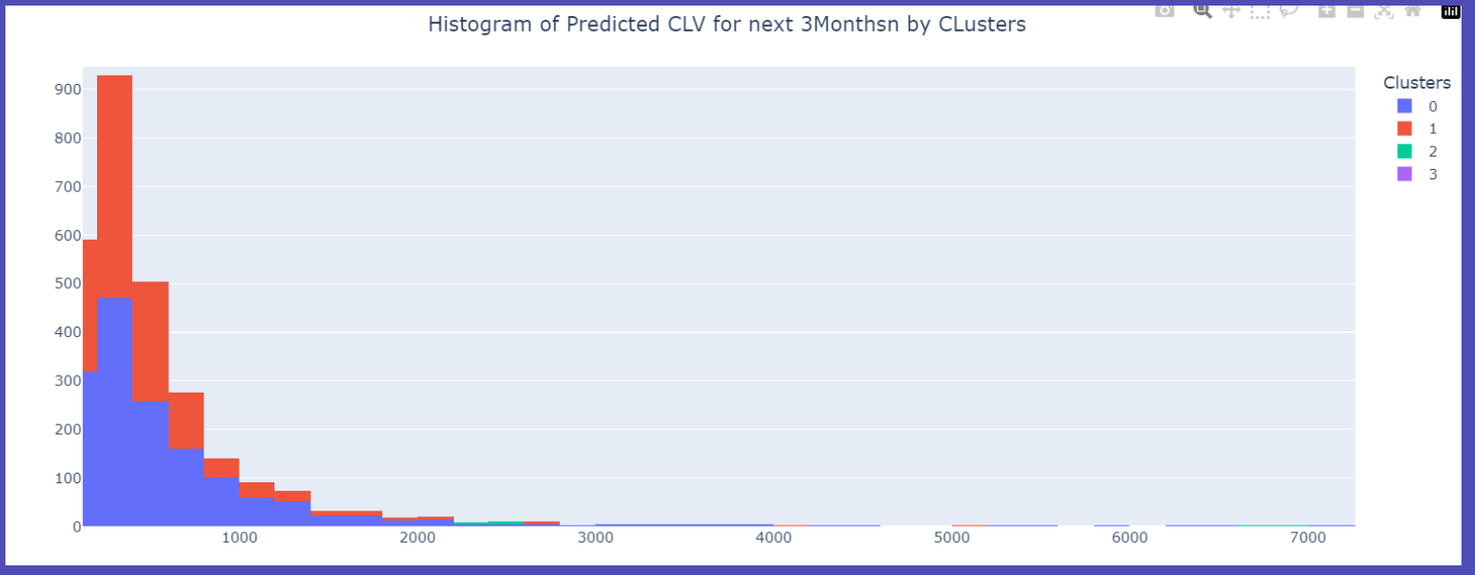
Customer ID	frequency	recency	T	monetary_value	Probability_alive	expected_purchases_in_90days	Expected_conditional_avg_revenue	Predict_CLV_3Months
12347	6.0	365.0	367.0	599.701667	0.999698	1.408736	569.977820	787.153453
12348	3.0	283.0	358.0	301.480000	0.999177	0.805907	333.785887	263.709035
12352	6.0	260.0	296.0	368.256667	0.999406	1.682277	376.175965	620.384166
12356	2.0	303.0	325.0	269.905000	0.999478	0.645368	324.041778	205.012619
12358	1.0	149.0	150.0	683.200000	0.999486	0.750390	539.904594	397.169912

Customer Segmentation using KMeans Algorithm

- Input features are Recency , Frequency , Lifespan(T), Monetary , Predicted CLV .
- Used Standard Scaling and PCA.

Kmeans Clustering





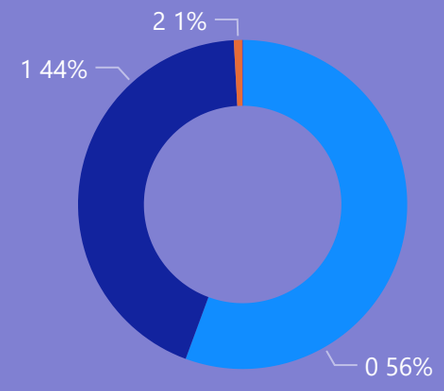
Cluster Analysis(For predicted CLV of Next 3months)

Clusters

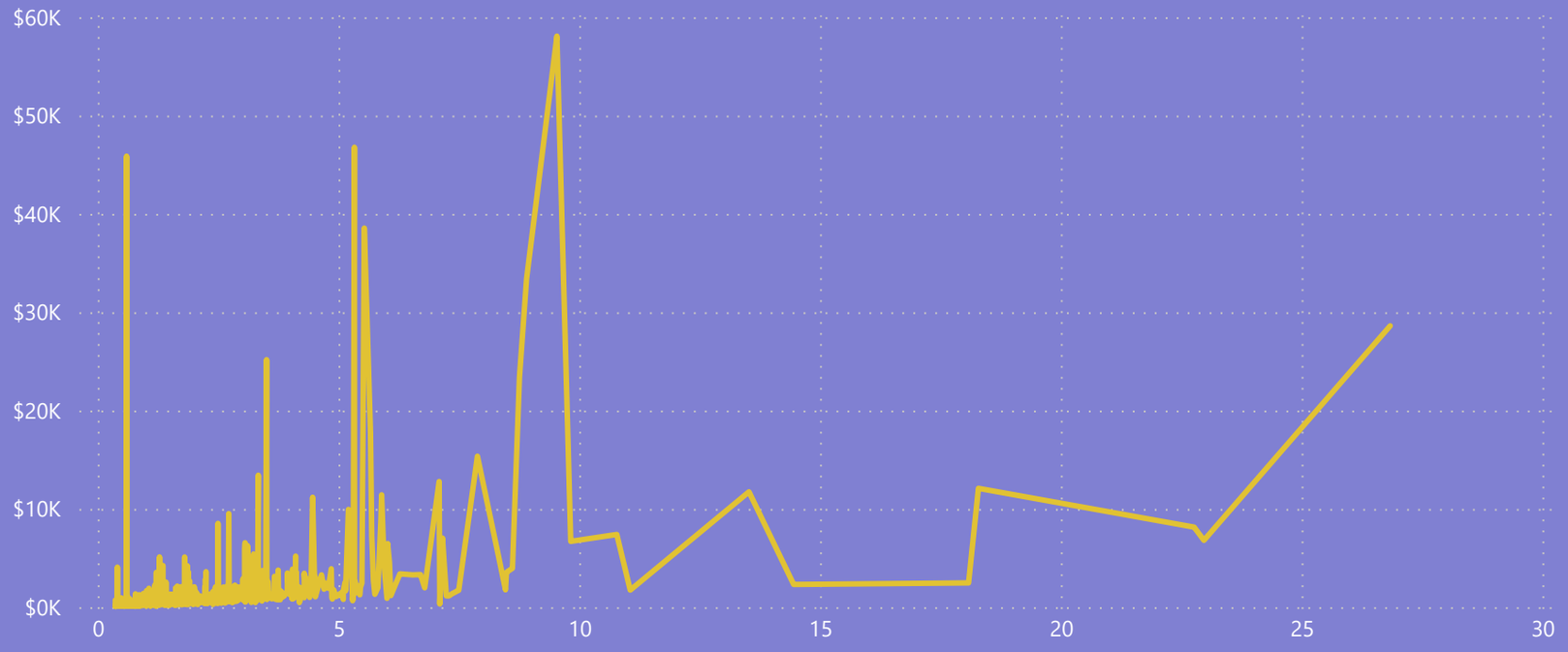
All

Clusters	Customers Count	Predicted Avg CLV	Predicted Avg No: Purchases	Past Avg Monetary	Predicted Max CLV
0	1552	$640	1.49	$416	$9,525
1	1214	$435	1.18	$375	$5,117
2	23	$17,934	10.63	$2,733	$58,083
3	1	$45,874	0.60	$1,68,470	$45,874

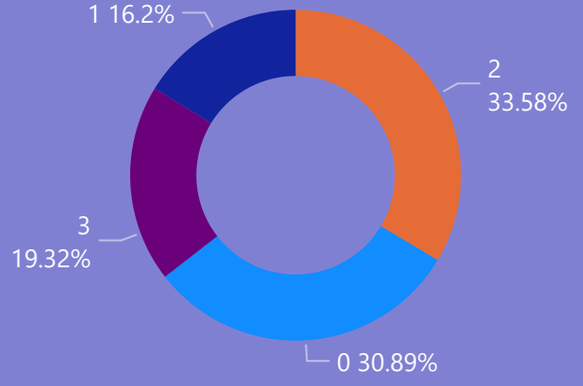
Number Of Customer by Clusters



Avg Predicted CLV vs Predicted Number of purchases



Avg Lifespan by Clusters



Thank You