Transactional Analysis of Insurance Data

Neha Savant

Table of Contents

Transactional Analysis of Insurance Data

1. Objective

To analyse patterns in transactions from a data-set of insurance policies that were sold to clients and to derive insights, inferences, and highlight anomalies, if any.

2. Data

The dataset contains 84524 policies. The unique policy holder base in this dataset is 77956. These policies were enforced from the May of 2017 to the March of 2018. Out of the 84K policies, 55% policies were paid between July and September 2018.

i. Assumptions

In order to club the data provided month wise into a single dataset, the payment status repeating over three months were shortened as follows-

- A policy that had at least one paid status over the period of 3 months was tagged 'Paid'.
- Policies that were paid with a grace status (less than 0.1% cases) were considered 'Paid'.
- Policies with missing Payment Status data or 'Lapsed' status in one or more files, were marked as 'Lapsed'

This is done taken into consideration that a payment update on policy would get recorded irrespective of the billing cycle since the months July to September fall in all the 4 billing cycles (Annual, Semi-Annually, Quarterly, Monthly).

ii. Handling Duplicates

With an attempted to analyse the data at a policy level, cases of duplicates were observed as below.

- **Multiple policy holders** – Only one owner per policy were retained. The record for the second owner was eliminated since all the other remained the same.
- **Change in Advisor Level** – Data recorded over three months had a change in advisor level of the agent owing to promotions in experience levels. In this case, the record with the highest advisor level was retained.
- **Change in Billing Type** – In this case, the record with the latest billing type (Direct/NACH) in three months was retained.
- Change in Manager – Over a period of three months, if the Agency manager ID changed, the record with the latest month was considered.

iii.   Additional Features

Channel Type – Sub Channels were largely grouped into channels –

- **Retail -** DSF Main, DM, APC
- **Corporate –** CAB, Banca, HDFC Salary Acct
- **Customer Driven –** D2C

No_of_policies_opted – This column was added to analyse cases of '1 Client: Multiple Policies'.

3. Analysis

i.   Is a non-payment of premium related to the annual premium amount?

Data shows that a higher premium guarantees more chances of payment. The cut-off of this amount is a premium of 18K.
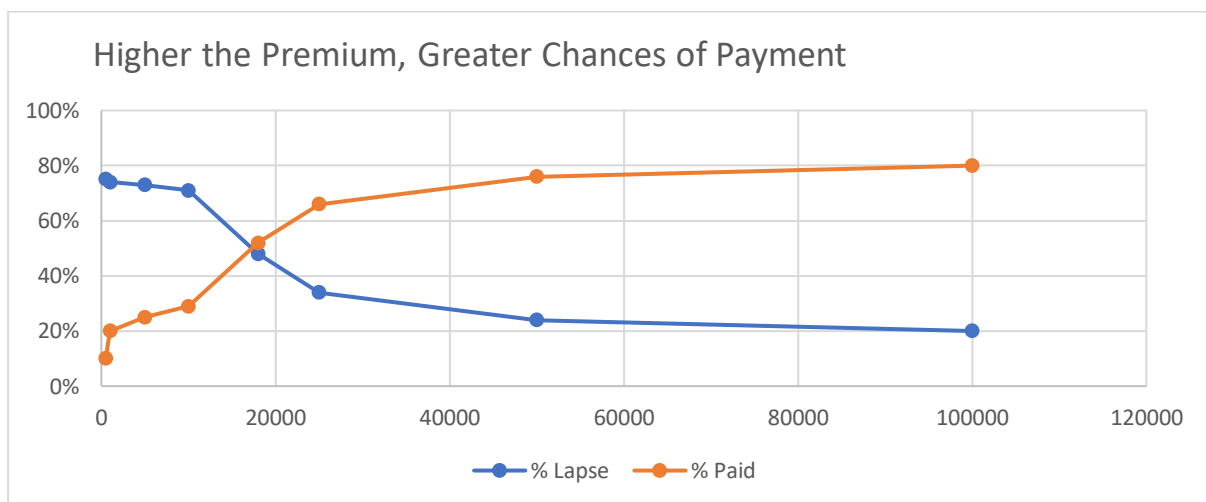

Figure 0.1

Higher premium involves options of the billing in parts or instalments either monthly, quarterly, or semi-annually. This can be attributed to promised payments as it is easy to pay over a period of several months than at one go.

ii.   (Non)-Profitable Products

About 10K policies of the product – BSLI Cancer Shield Plan were issued that year, but 80% of them did not receive a payment. It is therefore tagged as the most unprofitable plan so far.

Another unprofitable plan in terms of non-payment of the premium is the BSLI Hospital Plus Plan, but it is less impactful since only 682 policies were issued.

Policy	% Lapsed	% Paid	# Policies	Comment
BSLI Vision Endowment Plus	50%	50%	21158	
BSLI Cancer Shield Plan	80%	20%	9794	Most Unprofitable policy
BSLI Vision MoneyBack Plus	49%	51%	9233	
BSLI Income Assured	33%	67%	8866	Profitable policy
BSLI Wealth Secure	25%	75%	6758	Profitable policy
BSLI Vision Life Secure Plan	39%	61%	5522	
BSLI Vision LifeIncome Plan	38%	62%	4839	
BSLI Wealth Aspire	23%	77%	3407	Profitable policy
BSLI Vision Star	39%	61%	2582	
BSLI Secure Plus	25%	75%	2544	Profitable policy
BSLI Vision Endowment	45%	55%	1965	
BSLI Protector Plus 2014	42%	58%	1359	
BSLI Empower Pension Plan	38%	62%	1087	
BSLI Protect@Ease Plan	23%	77%	1048	Profitable policy
BSLI Savings Plan	57%	43%	923	
BSLI Fortune Elite	30%	70%	902	
BSLI Hospital Plus plan	88%	12%	682	Unprofitable policy
BSLI Easy Protect Plan	18%	82%	672	
BSLI Wealth Assure	18%	82%	514	
BSLI Protect @ Ease	18%	82%	339	
BSLI Future Guard	44%	56%	226	
BSLI Vision Regular Returns	46%	54%	175	
BSLI Guaranteed Future	49%	51%	45	
Bachat Child Plan	0%	100%	1	
Platinum Advantage Plan	0%	100%	1	

Table 0.1

On the other hand, policies with the highest rate of paid premium are from the categories –

'**Traditional Non-Par**' plans OR the **ULIP** plans as highlighted in green in the table below. This also highlights the fact that two separate set of customers take up these plans –

Either a steady return plan or a high-risk plan.

Product Categorization-

Product_Category	Segment	Ing_Product_Desc	Comment
Traditional	Non PAR	Bachat Child Plan	Fixed 5% return
Traditional	Non PAR	BSLI Cancer Shield Plan	
Traditional	Non PAR	BSLI Easy Protect Plan	
Traditional	Non PAR	BSLI Future Guard	
Traditional	Non PAR	BSLI Guaranteed Future	
Traditional	Non PAR	BSLI Hospital Plus plan	
Traditional	Non PAR	BSLI Income Assured	
Traditional	Non PAR	BSLI Protect @ Ease	
Traditional	Non PAR	BSLI Protect@Ease Plan	
Traditional	Non PAR	BSLI Protector Plus 2014	
Traditional	Non PAR	BSLI Secure Plus	
Traditional	PAR	BSLI Savings Plan	Returns based on company profits
Traditional	PAR	BSLI Vision Endowment	
Traditional	PAR	BSLI Vision Endowment Plus	
Traditional	PAR	BSLI Vision Life Secure Plan	
Traditional	PAR	BSLI Vision LifeIncome Plan	
Traditional	PAR	BSLI Vision MoneyBack Plus	
Traditional	PAR	BSLI Vision Regular Returns	
Traditional	PAR	BSLI Vision Star	
ULIP	Pension	BSLI Empower Pension Plan	Returns are market linked
ULIP	ULIP	BSLI Fortune Elite	
ULIP	ULIP	BSLI Wealth Aspire	
ULIP	ULIP	BSLI Wealth Assure	
ULIP	ULIP	BSLI Wealth Secure	
ULIP	ULIP	Platinum Advantage Plan	

Table 0.2

Wealth Plans – Data suggests that the 4 wealth plans below show the highest persistency (76% to 81%) in payment. These are BSLI Fortune Elite, BSLI Wealth Aspire, BSLI Wealth Assure, BSLI Wealth Secure. About 12K such policies were issued and have been profitable.

iii. Who's the influencer?

Data suggests that better the advisor level, higher is the chance that the policy holder pays the premium. Thus, policies shall be pushed through higher advisory level agents owing to their experience level, and helping a prospect understand the nuances of various plans and choosing the right policy.

Advisor Level	% Lapsed	% Paid
< LEV3	52%	44%
LEV3A	35%	65%
LEV4	28%	72%
LEV5	18%	82%
LEV5A	15%	85%
LEV10	18%	82%

Table 0.3

iv. Is it channel dependent or channel + product dependent?

It is observed that there are 81% chances of a premium being paid when the channel is customer driver (D2C), despite being only 482 cases in the dataset.

A major chunk is from the Retail channels and is a problematic area with only a 54% chance of paying the premium.

Channel	% Paid	#_Policies
Corporate	61%	9814
Customer Driven	82%	478
Retail	54%	**74350**

Table 0.4

Breaking down the retail channel –

- ULIP plans sold by agents with a > LEV4 advisory experience has 90% payment rate. As mentioned in point 5, it also means more experience to pitch in and make prospects understand about the right investment options.

- ULIP policies when sold in the West zone have an 83% payment rate while those when sold in the Northern belt have 75% payment rate. This indicates that the West and North zones of India are more open to riskier options than the rest of India. This is also a profitable belt as higher sum is insured in these belts, comparatively. The Sum Insured in ULIP policies (INR) are: North = 249 Mil, West = 223 Mil.

Zone	% Lapsed	% Paid
West 2	16%	84%
West 1	18%	82%
North 1	24%	76%
North 2	26%	74%
East 1	30%	70%
South	31%	69%
Central	33%	67%

Table 0.5

As a side note, ULIP Plans have greater returns and also have a higher payment rate, and is thus a WIN-WIN situation.

v.   Does the billing mode impact payment?

Though there are only 13% cases of NACH Payment as opposed to Direct ones, NACH guarantees 72% payment rate while that for Direct is a no show at 43%. Hence, if more policies are pushed to have NACH as the billing type, there is more than a $2/3^{rd}$ chance of regularity in premium payments.

vi.   Irregularity in payments – 1 Owner: Multiple Policies?

It is observed that there are cases where a single policy holder owns multiple policies. 107 such (policies) cases are seen in the **South zone**.

The payment behaviour of such policy holders is as below –

- The policy holder pays the first policy, gets issued several other after that but does not pay its premium. (1 policy holder: 1 agent: multiple policies)

Examples: Policy Holders– 2719435739, 9121247002, BH2450, 4980865001

- The policy holder buys multiple different policies but pays the last one with the highest premium amongst other. The back story could be that the policy holder may be reminded several times for payment and finally opts in for a sophisticated policy with a 'paid-up' value (generally lower than the expected price of the product)

Example – Policy Holder – 5525443545, 213622223, 840285850

- The policy holder buys same policy multiple times but pays for only one.

Example – Policy Holder – 5248453268, 5364566474

- Cases of 1:M policies are observed under the Regional Manager - Senthil Kumar.

4. Actionable

1. A policy with a premium of greater than 18K is more likely to be paid hence, pricing can be analysed.

2. ULIP and Wealth Plans ensure greater payment rate and should be pushed to the prospects first.

3. The Cancer Shield and Hospital Plus Policies though sell more, deeper customer analysis should be further carried out so that the prospect doesn't default on the premium at a later stage. Probably, the policy holder may have similar plans with other insurance corporations that have better benefits than BSLI.

4. If a policy buyer opts in for a second plan, check should be made if it is of the same kind, for example buying the same endowment policy twice. In that case, the policy holder is less likely to pay its premium.

5. Check should also be made that if the policy holder has paid the premium for an earlier policy before they opt in for a policy with a higher value. If they haven't paid the previous premium, there are higher chances of it not being paid in the future as well.

6. Cultivate/ educate more agents to LEV4 and above, impart trainings to existing agents

7. When the dataset was fit into random forest model, below were the important variables in determining the payment. Though the accuracy was only 68%, working on the dataset at the feature engineering level should boost the accuracy.

Variable Imp	
No_of_policies_opted	436.9767
Orginal_AP	3317.6325
Band	850.6308
AP_Band	875.1976
Product_Category	136.3490
Segment	307.8553
Ing_Product_Desc	1589.6862
Billing_Mode	359.9864
Billing_Type_V2	257.9529
Sub_Channel	279.0642
Channel_Type	104.2622
Zone_V1	1098.4517
Advisor_Level_V1	1594.8893
Issue_Month	582.2986
Inforce_Month	576.6237

Table 0.6

8. Adding information about the customer for example the age, salary bracket, background, medical history, family can help in an even better prediction and analysis of premium persistency.