LEAD SCORING CASE STUDY

PRESENTED BY

SHREYA SOLANKI
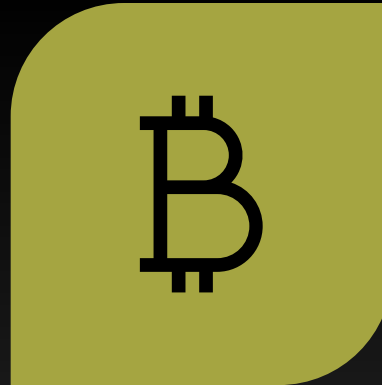
BATCH C10

PGDDS

Problem Statement

AN EDUCATION COMPANY NAMED X EDUCATION SELLS ONLINE COURSES TO INDUSTRY PROFESSIONALS.

NOW, ALTHOUGH X EDUCATION GETS A LOT OF LEADS, ITS LEAD CONVERSION RATE IS VERY POOR OF ABOUT 30%.

THE COMPANY WANTS TO INCREASE IT TO 80%

Goal

Build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads.

A higher score would mean that the lead is hot, i.e. is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted.

TOP FACTORS THAT IMPACT THE CONVERSION OF LEADS

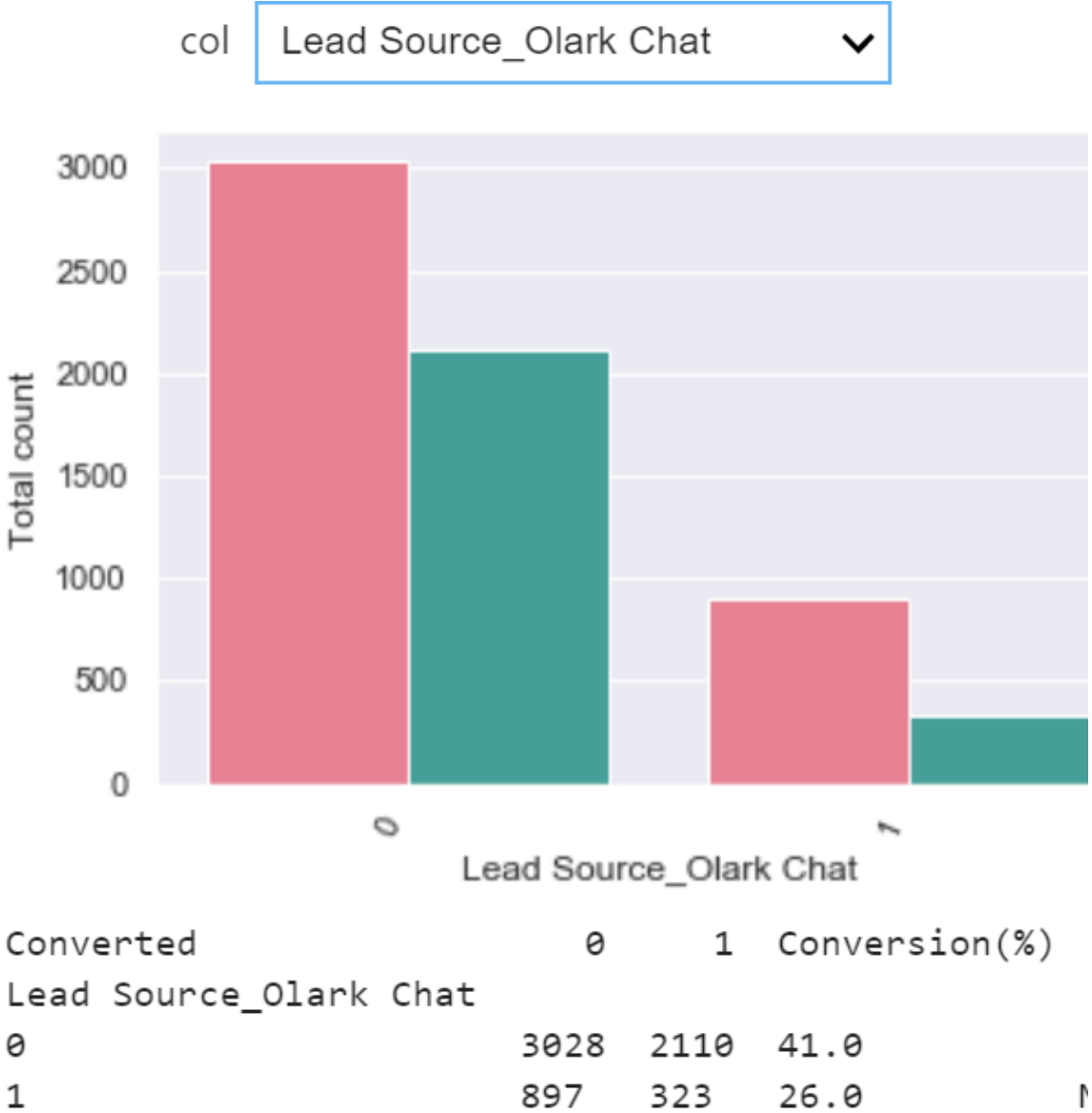| Features |
|---|
| Tags_Will revert after reading the email |
| Total Time Spent on Website |
| TotalVisits |
| Lead Origin_Lead Add Form |
| Last Notable Activity_SMS Sent |
| Last Notable Activity_Modified |
| Lead Source_Olark Chat |
| Lead Profile_Potential Lead |
| Lead Source_Welingak Website |
| Tags_Closed by Horizzon |
| Lead Quality_Not Sure |
| Do Not Email_Yes |
| Tags_Lost to EINS |
| Lead Profile_Other Leads |
| Last Notable Activity_Olark Chat Conversation |

LETS SEE HOW THESE FACTORS ARE IMPACTFUL

WHAT MEASURES SHOULD BE TAKEN TO INCREASE THE CONVERSION RATE OF HOT LEADS?

Lead Source : Olark Chat

This means that there is a lot of traffic on this Source

Increasing the number of representatives who could guide the prospects so that they are not missed out and hence increase the number of HOTS


```
Converted                        0     1   Conversion(%)
Lead Source_Olark Chat
0                             3028  2110   41.0
1                              897   323   26.0
```
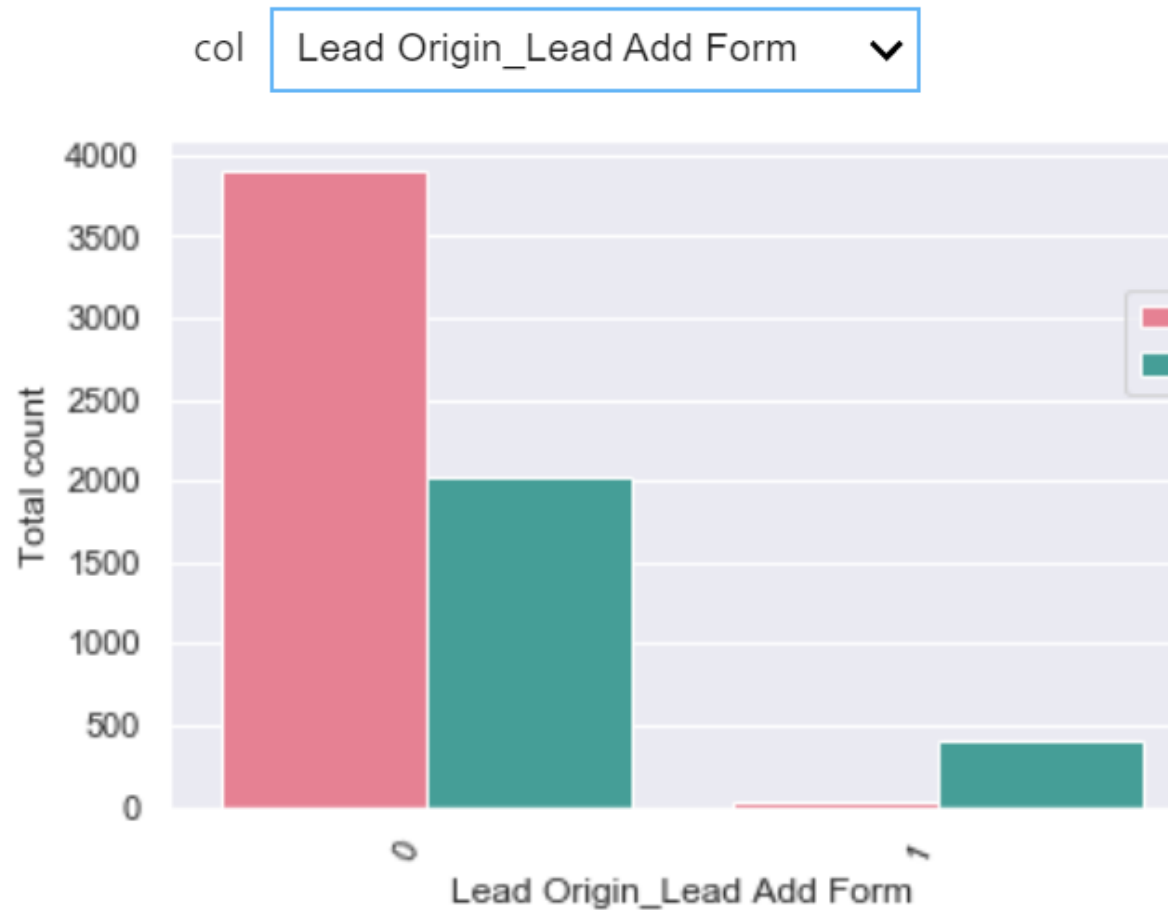
Lead Source Welingkak

Whichever prospective leads are using this website have successfully converted

Time to make this a hit and try to get more leads through the website by promoting this one


```
col   Lead Source_Welingak Website ⌄
```
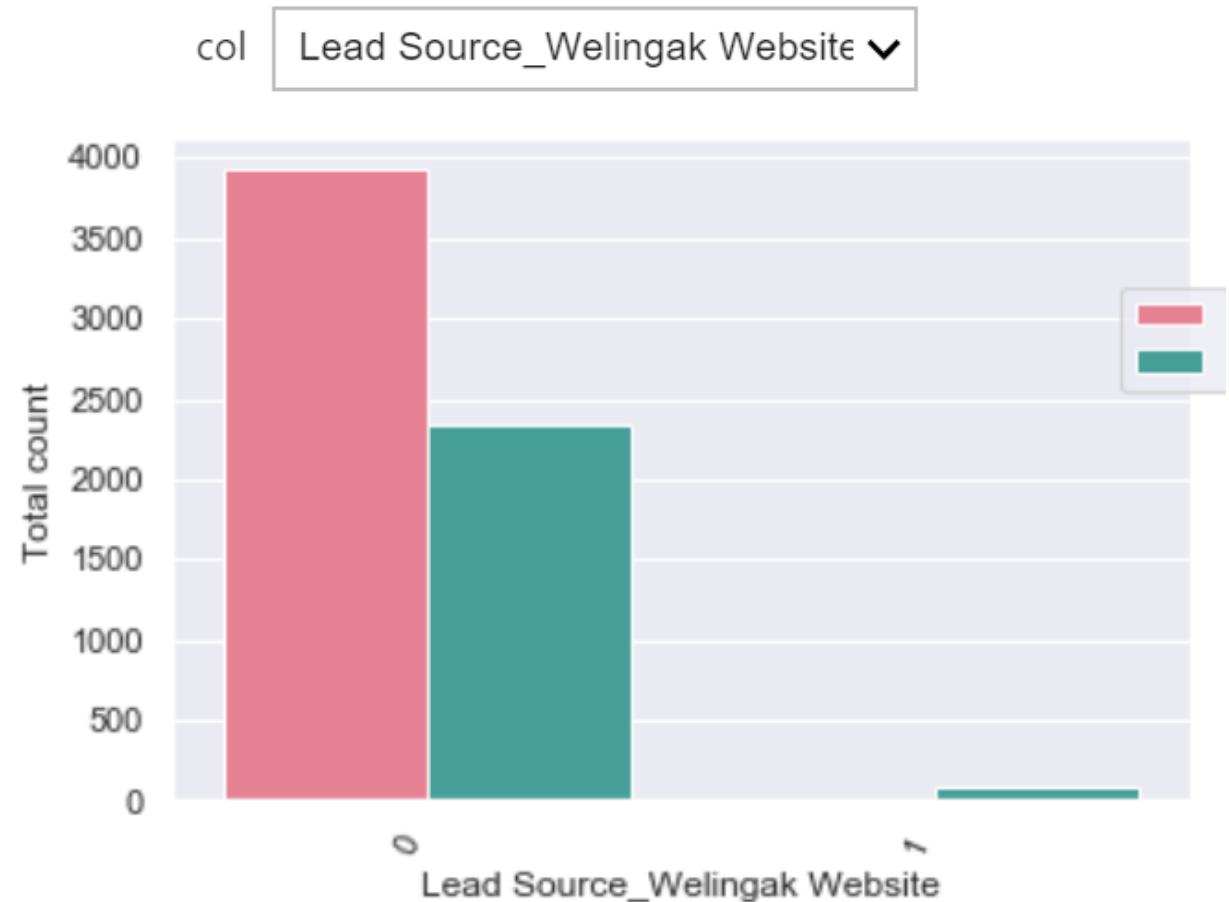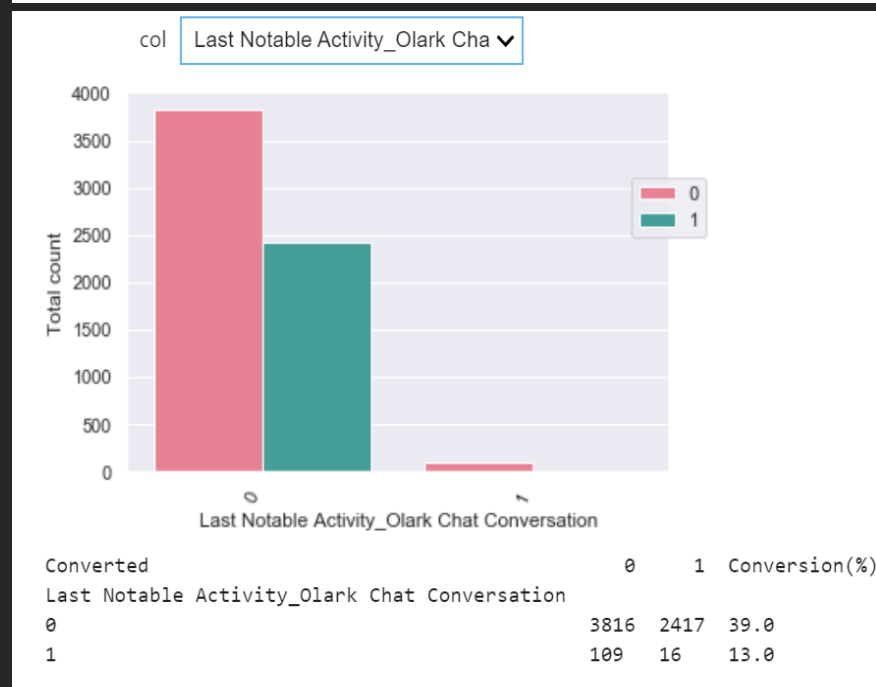
```
Converted                              0     1   Conversion(%)
Lead Source_Welingak Website
0                                   3923  2340    37.0
1                                      2    93    98.0
```

LEAD ADD FORM HAS A CONVERSION RATE OF 93%

MOST FOCUS SHOULD BE GIVEN TO THIS ORIGIN
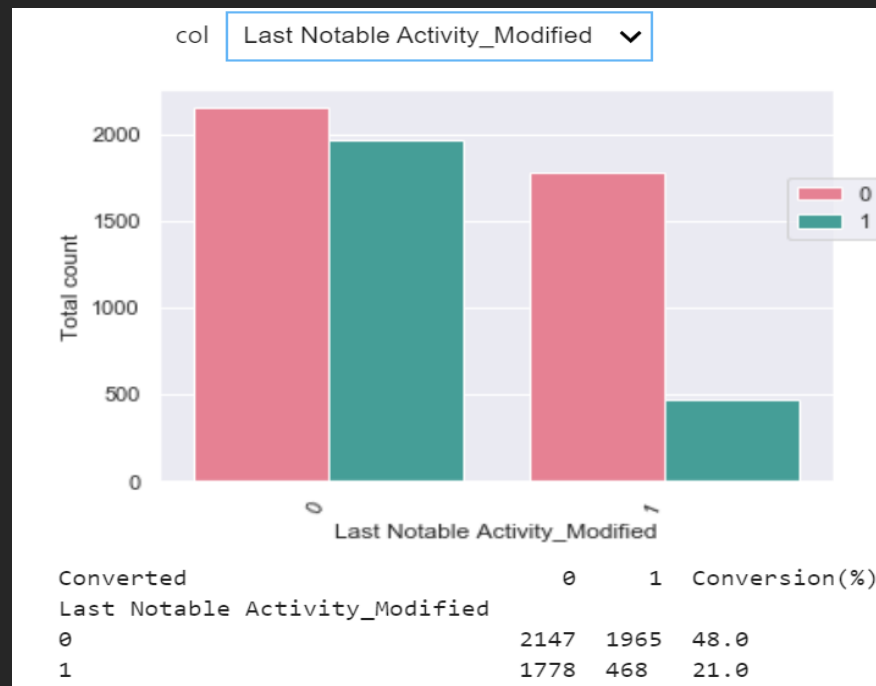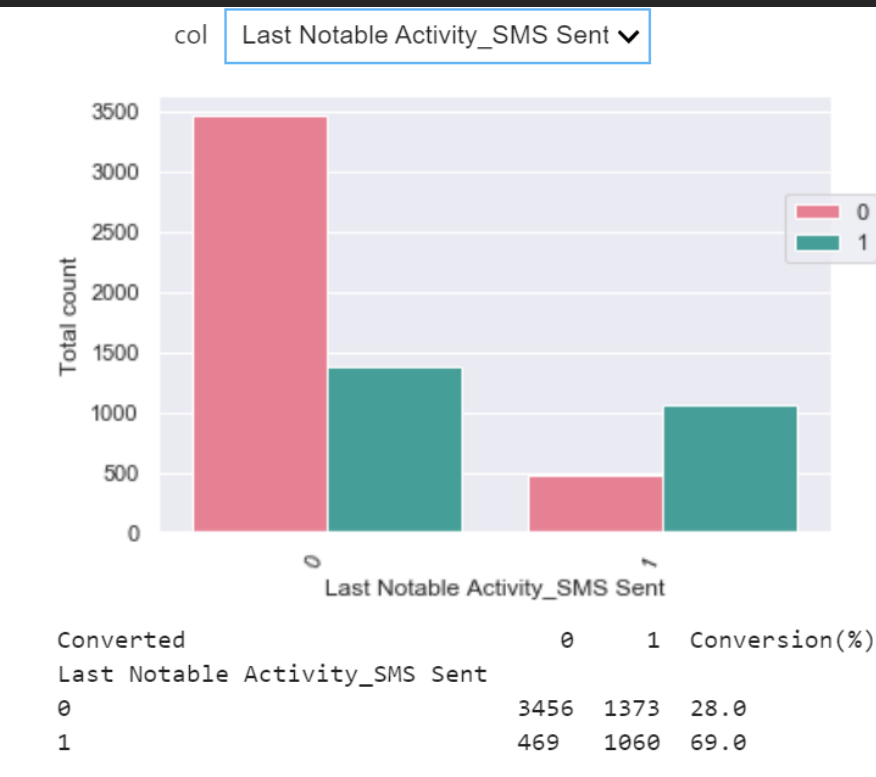
All these Last Activities of the users are those which should be focused on more as compared to other activities, Highest conversion rate is in SMS sent activity. Keeping track of this one would be beneficial

Potential leads:
Leads whose profile is mentioned as Potential, are indeed having the potential of converting to Hot leads.
There is a 78% conversion rate of this category.
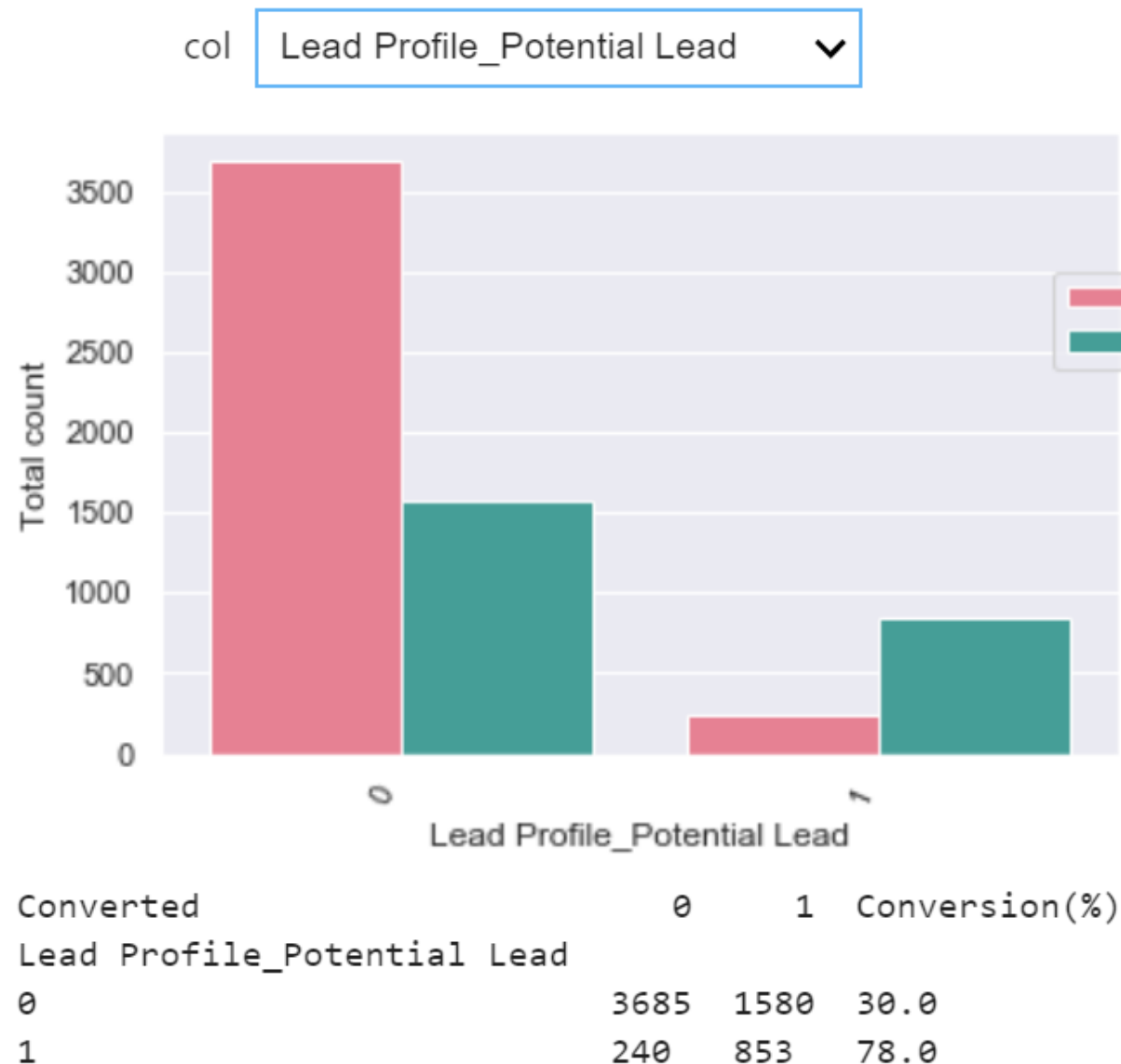This lead profile should be taken specific care when guiding them, as they have a high probability of conversion

col Lead Profile_Potential Lead

Lead Profile_Potential Lead

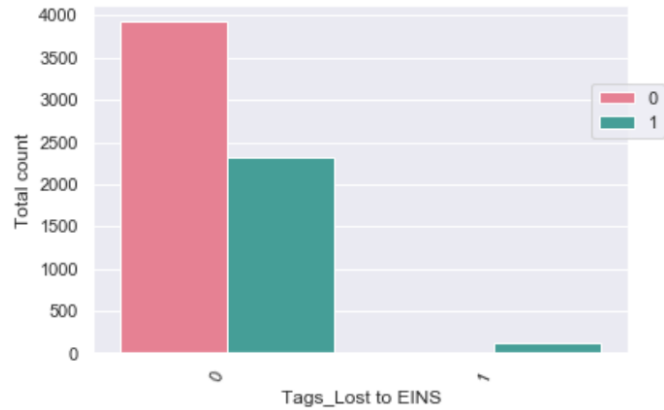| Converted	0	1	Conversion(%)
Lead Profile_Potential Lead			
0	3685	1580	30.0
1	240	853	78.0

Tags :

Lost to EINS

Closed by Horizon

Have a very high conversion rate, highly correlated to Conversion

Keeping an eye on these tags is important for conversion rate

col Tags_Lost to EINS ⌄

Converted   0   1   Conversion(%)
Tags_Lost to EINS
0   3923   2313   37.0
1   2   120   98.0   None

col Tags_Will revert after reading th ⌄

Converted   0   1   Conversion(%)
Tags_Will revert after reading the email
1   1771 1977   53.0
0   2154   456   17.0

col Tags_Closed by Horizzon ⌄

Converted   0   1   Conversion(%)
Tags_Closed by Horizzon
0   3923   2221   36.0
1   2   212   99.0

TO ACHIEVE A CONVERSION RATE OF AROUND 80%:

--NEED TO NOT MISS THOSE PROSPECTS WHICH CAN TURN INTO HOT LEADS WHICH IS RECALL

--SHOULD NOT OVERESTIMATE A COLD LEAD WHICH IS PRECISION

A right choice of Probability at which the Lead should be considered as potential and can turn into a hot lead is the major need:

Hence using this model we achieve a Right trade off point for having 0.38 i.e. 38% probability of a Prospect is good enough to target as HOT lead and work more on the strategies of that lead rather than Cold ones.
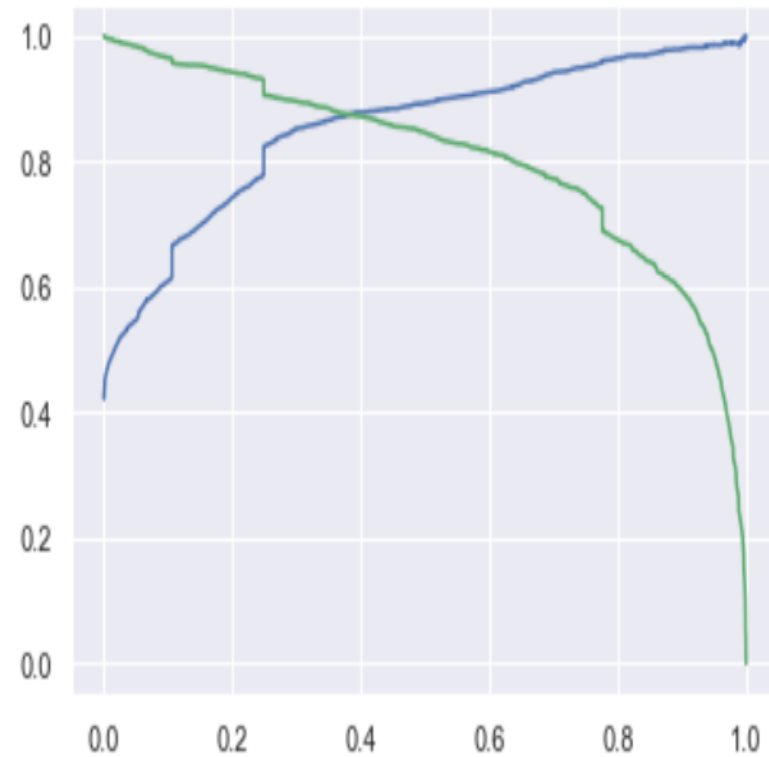
0.38 is the tradeoff between Precision and Recall -

thus we can safely choose to consider any Prospect Lead with Conversion Probability higher than 38 % to be a hot Lead

Lead Prediction on the basis of probability set by the model

Final_pred column shows the Conversion Probablity of a Prospect Lead

Lead_Score above 38 shows a potential lead and Has the capability to turn into HOT lead.

Lead_Score 38 mathematically specifies that 38% probability is good to go for targettting the lead in order to achieve 80% Conversion

	LeadID	Converted	Conv_Prob	final_pred	Lead_Score
0	1786	1	0.776454	1	78
1	2469	0	0.066023	0	7
2	4177	0	0.013659	0	1
3	3912	1	0.879599	1	88
4	5138	0	0.776454	1	78
5	7202	1	0.998769	1	100
6	6678	0	0.025713	0	3
7	284	0	0.590213	1	59
8	2741	0	0.107175	0	11
9	6790	0	0.005487	0	1
10	770	0	0.775202	1	78
11	1979	1	0.983942	1	98
12	2706	0	0.719485	1	72
13	2354	1	0.999191	1	100

14	3643	1	0.947821	1	95
15	7973	0	0.000386	0	0
16	6041	0	0.074779	0	7
17	6351	0	0.139613	0	14
18	1911	0	0.152152	0	15
19	5068	1	0.964977	1	96
20	5163	0	0.042224	0	4
21	1560	1	0.491074	1	49
22	7792	0	0.000911	0	0
23	7242	0	0.084295	0	8
24	5611	1	0.766236	1	77
25	1053	1	0.995206	1	100
26	5077	1	0.935720	1	94
27	9049	1	0.934988	1	93

How 80% conversion is achieved?

Even a Lead that was not converted as per the available history, those leads could be guided properly based on the Lead_Score provided, without wasting any Resources on Low Probability lead

Sensitivity is the metric to show that the 80% conversion is achieved on the data .

Any new Prospect lead , generates a score above 38 in the model, then , the company needs to take measures to guide the lead.

```
Sensitivity of the model on test data is  0.88
```