

# Subjective questions

---

# 1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Based on the coefficient values from below screenshot, the following are the top three variables that contribute most towards the probability of a lead getting converted :

- Total Time Spent on Website
- Lead Add Form (from Lead Origin)
- Had a Phone Conversation ( from Last Notable Activity)

	coef	std err	z	P> z	[0.025	0.975]
const	-2.2778	0.108	-21.100	0.000	-2.489	-2.066
Do Not Email	-1.1622	0.171	-6.798	0.000	-1.497	-0.827
TotalVisits	1.1050	0.248	4.459	0.000	0.619	1.591
Total Time Spent on Website	4.5754	0.171	26.798	0.000	4.241	4.910
LeadOrigin_API	0.3591	0.092	3.908	0.000	0.179	0.539
LeadOrigin_Lead Add Form	3.9446	0.224	17.580	0.000	3.505	4.384
LeadSource_Olark Chat	1.3759	0.135	10.178	0.000	1.111	1.641
LeadSource_Welingak Website	2.0688	0.744	2.782	0.005	0.611	3.526
LastActivity_Not Sure	-1.2383	0.457	-2.708	0.007	-2.135	-0.342
LastActivity_Olark Chat Conversation	-1.0035	0.171	-5.874	0.000	-1.338	-0.669
LastActivity_SMS Sent	1.2453	0.078	16.027	0.000	1.093	1.398
CurrentOccupation_No Information	-1.2063	0.089	-13.551	0.000	-1.381	-1.032
CurrentOccupation_Working Professional	2.3632	0.182	12.971	0.000	2.006	2.720
LastNotableActivity_Had a Phone Conversation	3.3426	1.133	2.949	0.003	1.121	5.564
LastNotableActivity_Modified	-0.7868	0.084	-9.393	0.000	-0.951	-0.623
LastNotableActivity_Page Visited on Website	-0.5113	0.213	-2.399	0.016	-0.929	-0.094
LastNotableActivity_Unreachable	1.6170	0.555	2.911	0.004	0.528	2.706

2. What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

---

Based on the coefficient values from the screen shot in the question above, the following are the top three categorical/dummy variables that should be focused the most in order to increase the probability of lead conversion :

- Lead Add Form (from Lead Origin)
- Had a Phone Conversation ( from Last Notable Activity)
- Working Professional ( from What is your current occupation)

3. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers who have been predicted as 1 by the model) to be converted and hence, want to make phone calls to as much of such people as possible. Suggest a good strategy they should employ at this stage.

In the below image, the final prediction is calculated based on a optimal cut off value of 0.37.

In order to make the sales aggressive, the company may contact all the leads which have a conversion probability (value = 1) under a cut off 0.3 (column 0.3 highlighted in yellow).

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
0	0	0.061044	5493		0	1	0	0	0	0	0	0	0	0	0	6
1	0	0.090376	8064		0	1	0	0	0	0	0	0	0	0	0	9
2	0	0.028539	4716		0	1	0	0	0	0	0	0	0	0	0	3
3	0	0.367527	9117		0	1	1	1	1	0	0	0	0	0	0	37
4	1	0.470169	2402		0	1	1	1	1	1	0	0	0	0	1	47
5	0	0.028210	1796		0	1	0	0	0	0	0	0	0	0	0	3
6	0	0.028210	1120		0	1	0	0	0	0	0	0	0	0	0	3
7	0	0.051276	253		0	1	0	0	0	0	0	0	0	0	0	5
8	0	0.122101	1491		0	1	1	0	0	0	0	0	0	0	0	12
9	1	0.339668	2004		0	1	1	1	1	0	0	0	0	0	0	34
10	0	0.248904	1792		0	1	1	1	0	0	0	0	0	0	0	25
11	0	0.088416	1944		0	1	0	0	0	0	0	0	0	0	0	9
12	0	0.033781	3879		0	1	0	0	0	0	0	0	0	0	0	3
13	0	0.269670	4084		0	1	1	1	0	0	0	0	0	0	0	27
14	0	0.111327	1338		0	1	1	0	0	0	0	0	0	0	0	11
15	1	0.841140	4371		1	1	1	1	1	1	1	1	1	1	0	84
16	0	0.163065	6941		0	1	1	0	0	0	0	0	0	0	0	16
17	0	0.211684	6600		0	1	1	1	0	0	0	0	0	0	0	21
18	0	0.013656	6818		0	1	0	0	0	0	0	0	0	0	0	1
19	0	0.203079	5930		0	1	1	1	0	0	0	0	0	0	0	20

4. Similarly, at times, the company reaches its target for a quarter before the deadline. During this time, the company wants the sales team to focus on some new work as well. So during this time, the company's aim is to not make phone calls unless it's extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1 highlighted in yellow color) under column 0.7.

However, the flipside here would be that, we may miss out on those leads that are actually converted but then the model wrongly predicted them as not converted.

(See red highlights in the image below). This should not be a major cause for concern as the target has already been achieved.

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
0	0	0.061044	5493		0	1	0	0	0	0	0	0	0	0	0	6
1	0	0.090376	8064		0	1	0	0	0	0	0	0	0	0	0	9
2	0	0.028539	4716		0	1	0	0	0	0	0	0	0	0	0	3
3	0	0.367527	9117		0	1	1	1	1	0	0	0	0	0	0	37
4	1	0.470169	2402		0	1	1	1	1	1	0	0	0	0	1	47
5	0	0.028210	1796		0	1	0	0	0	0	0	0	0	0	0	3
6	0	0.028210	1120		0	1	0	0	0	0	0	0	0	0	0	3
7	0	0.051276	253		0	1	0	0	0	0	0	0	0	0	0	5
8	0	0.122101	1491		0	1	1	0	0	0	0	0	0	0	0	12
9	1	0.339668	2004		0	1	1	1	1	0	0	0	0	0	0	34
10	0	0.248904	1792		0	1	1	1	0	0	0	0	0	0	0	25
11	0	0.088416	1944		0	1	0	0	0	0	0	0	0	0	0	9
12	0	0.033781	3879		0	1	0	0	0	0	0	0	0	0	0	3
13	0	0.269670	4084		0	1	1	1	0	0	0	0	0	0	0	27
14	0	0.111327	1338		0	1	1	0	0	0	0	0	0	0	0	11
15	1	0.841140	4371		1	1	1	1	1	1	1	1	1	0	1	84
16	0	0.163065	6941		0	1	1	0	0	0	0	0	0	0	0	16
17	0	0.211684	6600		0	1	1	1	0	0	0	0	0	0	0	21
18	0	0.013656	6818		0	1	0	0	0	0	0	0	0	0	0	1
19	0	0.203079	5930		0	1	1	1	0	0	0	0	0	0	0	20