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

   Solution:

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

   i. Total Time Spent on Website
   ii. Lead Add Form (from Lead Origin)
   iii. Working Professional (from Current Occupation)

	coef
const	-0.8058
Do Not Email	-1.3065
TotalVisits	0.9369
Total Time Spent on Website	4.5689
LeadOrigin_Lead Add Form	3.8169
LeadSource_Olark Chat	1.5488
LeadSource_Welingak Website	2.1293
LastActivity_Converted to Lead	-0.8189
LastActivity_Email Bounced	-1.1175
LastActivity_No Available	-1.6980
LastActivity_Olark Chat Conversation	-1.2655
CurrentOccupation_No Information	-1.1329
CurrentOccupation_Working Professional	2.3504
LastNotableActivity_Email Link Clicked	-1.6310
LastNotableActivity_Email Opened	-1.3862
LastNotableActivity_Modified	-1.6574
LastNotableActivity_Olark Chat Conversation	-1.3982
LastNotableActivity_Page Visited on Website	-1.8224

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?**

   Solution:

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

i. Lead Add Form (from Lead Origin)
ii. Working Professional (from Current Occupation)
iii. Welingak Website (from Lead Source)

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.**

Solution:

In the below image, the final prediction is calculated based on an optimal cutoff value of 0.36.
In order to make the sales aggressive, the company may contact all the leads which have a conversion probability (value = 1) under a lower cutoff (**Lower than 0.36**). This is because, the number of leads suggested by the model increase as the cutoff decreases. The cutoff should be based on the requirement of the company. For example, if the cutoff is **0.3** the case will be as shown below (it is underlined in yellow)

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	...	0.32	0.33	0.34	0.35	0.36	0.37	0.38	0.39	final_predicted	lead_score
0	0	0.104052	5493	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	10
1	0	0.109247	8064	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	11
2	0	0.013730	4716	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
3	0	0.344528	9117	0	1	1	1	1	0	0	...	1	1	1	0	0	0	0	0	0	34
4	1	0.475239	2402	0	1	1	1	1	1	0	...	1	1	1	1	1	1	1	1	1	48
5	0	0.035135	1796	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	4
6	0	0.035135	1120	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	4
7	0	0.057269	253	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	6
8	0	0.125597	1491	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	13
9	1	0.387121	2004	0	1	1	1	1	0	0	...	1	1	1	1	1	1	1	0	1	39
10	0	0.356852	1792	0	1	1	1	1	0	0	...	1	1	1	1	0	0	0	0	0	36
11	0	0.101568	1944	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	10
12	0	0.062934	3879	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	6
13	0	0.269818	4084	0	1	1	1	0	0	0	...	0	0	0	0	0	0	0	0	0	27
14	0	0.117897	1338	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	12
15	1	0.835471	4371	1	1	1	1	1	1	1	...	1	1	1	1	1	1	1	1	1	84

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.**

Solution:

In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1) under a higher cutoff (**Higher than 0.36**). This is because, the number of leads suggested by the model decrease as the cutoff increases. The cutoff should be based on the requirement of the company. For example, if the cutoff is **0.5** the case will be as shown below (it is highlighted in yellow) However, the flipside is we may miss those leads that are actually converted. This should not be a problem as th Target has been already achieved.

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	...	0.32	0.33	0.34	0.35	0.36	0.37	0.38	0.39	final_predicted	lead_score
0	0	0.104052	5493	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	10
1	0	0.109247	8064	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	11
2	0	0.013730	4716	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
3	0	0.344528	9117	0	1	1	1	1	0	0	...	1	1	1	0	0	0	0	0	0	34
4	1	0.475239	2402	0	1	1	1	1	1	0	...	1	1	1	1	1	1	1	1	1	48
5	0	0.035135	1796	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	4
6	0	0.035135	1120	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	4
7	0	0.057269	253	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	6
8	0	0.125597	1491	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	13
9	1	0.387121	2004	0	1	1	1	1	0	0	...	1	1	1	1	1	1	1	0	1	39
10	0	0.356852	1792	0	1	1	1	1	0	0	...	1	1	1	1	0	0	0	0	0	36
11	0	0.101568	1944	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	10
12	0	0.062934	3879	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	6
13	0	0.269818	4084	0	1	1	1	0	0	0	...	0	0	0	0	0	0	0	0	0	27
14	0	0.117897	1338	0	1	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	12
15	1	0.835471	4371	1	1	1	1	1	1	1	...	1	1	1	1	1	1	1	1	1	84