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

   Solution:

   According to the coefficient values displayed in the screenshot below, the top three variables that have the greatest influence on the likelihood of converting a lead are as follows:

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

	coef
const	-2.3134
Do Not Email	-1.1067
TotalVisits	1.5038
Total Time Spent on Website	4.5275
LeadOrigin_Landing Page Submission	-0.3253
LeadOrigin_Lead Add Form	3.7313
LeadSource_Olark Chat	1.5706
LeadSource_Welingak Website	2.1406
LastActivity_Email Bounced	-0.9859
LastActivity_Not Sure	-1.5943
LastActivity_Olark Chat Conversation	-1.3427
LastActivity_SMS Sent	1.3544
CurrentOccupation_No Information	-1.3041
CurrentOccupation_Working Professional	2.6421
LastNotableActivity_Had a Phone Conversation	3.2904
LastNotableActivity_Unreachable	1.7567

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:

   Once more, considering the coefficient values depicted in the aforementioned screenshot, the following are the three primary categorical/dummy variables that warrant the most attention for enhancing the likelihood of lead conversion:

   a) Lead Add Form (from Lead Origin)
   b) Had a Phone Conversation (from Last Notable Activity)
   c) 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.

   Solution:

   In the image provided below, the ultimate prediction is determined using an optimal cutoff value of 0.37.

   To intensify sales efforts, the company can reach out to all leads with a conversion probability of 1 that falls below a cutoff value of 0.3 (highlighted in yellow in the column labeled 0.3).

	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	0.600894	2240	1	1	1	1	1	1	1	1	0	0	0	1	60
1	0	0.794876	113	1	1	1	1	1	1	1	1	1	0	0	1	79
2	1	0.802664	4132	1	1	1	1	1	1	1	1	1	1	0	1	80
3	0	0.089476	5573	0	1	0	0	0	0	0	0	0	0	0	0	9
4	0	0.029809	1109	0	1	0	0	0	0	0	0	0	0	0	0	3
5	0	0.041753	2282	0	1	0	0	0	0	0	0	0	0	0	0	4
6	1	0.767093	2976	1	1	1	1	1	1	1	1	1	0	0	1	77
7	0	0.418665	8431	0	1	1	1	1	1	0	0	0	0	0	1	42
8	1	0.805010	2770	1	1	1	1	1	1	1	1	1	1	0	1	81
9	1	0.995568	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.950498	2943	1	1	1	1	1	1	1	1	1	1	1	1	95
11	0	0.197985	1196	0	1	1	0	0	0	0	0	0	0	0	0	20
12	1	0.648304	8874	1	1	1	1	1	1	1	1	0	0	0	1	65
13	0	0.106911	1491	0	1	1	0	0	0	0	0	0	0	0	0	11
14	0	0.110507	7676	0	1	1	0	0	0	0	0	0	0	0	0	11
15	1	0.490611	8750	0	1	1	1	1	1	0	0	0	0	0	1	49
16	1	0.805010	5049	1	1	1	1	1	1	1	1	1	1	0	1	81
17	0	0.648304	5691	1	1	1	1	1	1	1	1	0	0	0	1	65
18	1	0.380602	5773	0	1	1	1	1	0	0	0	0	0	0	1	38
19	0	0.032619	3906	0	1	0	0	0	0	0	0	0	0	0	0	3

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:

To reduce the number of unproductive phone calls, the company can reach out to all leads highlighted in yellow (with a conversion probability value of 1) in the column labeled 0.7. However, a potential drawback is that we might overlook leads that were actually converted but were inaccurately predicted as non-converted by the model (indicated by red highlights in the image below). This shouldn't be a significant concern since 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	0.600894	2240	1	1	1	1	1	1	1	1	0	0	0	1	60
1	0	0.794876	113	1	1	1	1	1	1	1	1	1	0	0	1	79
2	1	0.802664	4132	1	1	1	1	1	1	1	1	1	1	0	1	80
3	0	0.089476	5573	0	1	0	0	0	0	0	0	0	0	0	0	9
4	0	0.029809	1109	0	1	0	0	0	0	0	0	0	0	0	0	3
5	0	0.041753	2282	0	1	0	0	0	0	0	0	0	0	0	0	4
6	1	0.767093	2976	1	1	1	1	1	1	1	1	1	0	0	1	77
7	0	0.418665	8431	0	1	1	1	1	1	0	0	0	0	0	1	42
8	1	0.805010	2770	1	1	1	1	1	1	1	1	1	1	0	1	81
9	1	0.995568	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.950498	2943	1	1	1	1	1	1	1	1	1	1	1	1	95
11	0	0.197985	1196	0	1	1	0	0	0	0	0	0	0	0	0	20
12	1	0.648304	8874	1	1	1	1	1	1	1	1	0	0	0	1	65
13	0	0.106911	1491	0	1	1	0	0	0	0	0	0	0	0	0	11
14	0	0.110507	7676	0	1	1	0	0	0	0	0	0	0	0	0	11
15	1	0.490611	8750	0	1	1	1	1	1	0	0	0	0	0	1	49
16	1	0.805010	5049	1	1	1	1	1	1	1	1	1	1	0	1	81
17	0	0.648304	5691	1	1	1	1	1	1	1	1	0	0	0	1	65
18	1	0.380602	5773	0	1	1	1	1	0	0	0	0	0	0	1	38
19	0	0.032619	3906	0	1	0	0	0	0	0	0	0	0	0	0	3