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 below screenshot, the following are the top three variables that contribute most towards the probability of a lead getting converted:
   a) Total Time Spent on Website
   b) Lead Add Form (from Lead Origin)
   c) Had a Phone Conversation (from Last Notable Activity)

	Coef.
const	-2.5028
Do Not Email	-1.1452
TotalVisits	1.7397
Total Time Spent on Website	4.5109
Page Views Per Visit	-0.8591
LeadOrigin_Lead Add Form	3.6099
LeadSource_Olark Chat	1.4653
LeadSource_Welingak Website	2.0660
LastActivity_Email Opened	0.5006
LastActivity_Olark Chat Conversation	-0.6438
LastActivity_SMS Sent	1.6955
CurrentOccupation_No Information	-1.2447
CurrentOccupation_Working Professional	2.6123
LastNotableActivity_Had a Phone Conversation	3.5184
LastNotableActivity_Modified	-0.5330
LastNotableActivity_Unreachable	2.0055

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:
   Again, 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 :
   We can focus on already existing important features, and also we can concentrating on low performing features and improve on them.
   a) Do Not Email
   b) CurrentOccupation_No Information
   c) Pages Views Per Visit

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 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 inyellow).

	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	0.692600	2240	1	1	1	1	1	1	1	1	0	0	0	1	69
1	0	0.539248	113	1	1	1	1	1	1	1	0	0	0	0	1	54
2	1	0.718698	4132	1	1	1	1	1	1	1	1	1	0	0	1	72
3	0	0.133628	5573	0	1	1	0	0	0	0	0	0	0	0	0	13
4	0	0.014225	1109	0	1	0	0	0	0	0	0	0	0	0	0	1
5	0	0.031692	2282	0	1	0	0	0	0	0	0	0	0	0	0	3
6	1	0.854208	2976	1	1	1	1	1	1	1	1	1	1	0	1	85
7	0	0.399257	8431	0	1	1	1	1	0	0	0	0	0	0	1	40
8	1	0.745493	2770	1	1	1	1	1	1	1	1	1	0	0	1	75
9	1	0.995570	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.957687	2943	1	1	1	1	1	1	1	1	1	1	1	1	96
11	0	0.243037	1196	0	1	1	1	0	0	0	0	0	0	0	0	24
12	1	0.531214	8874	1	1	1	1	1	1	1	0	0	0	0	1	53
13	0	0.130486	1491	0	1	1	0	0	0	0	0	0	0	0	0	13
14	0	0.098482	7676	0	1	0	0	0	0	0	0	0	0	0	0	10
15	1	0.460246	8750	0	1	1	1	1	1	0	0	0	0	0	1	46
16	1	0.833093	5049	1	1	1	1	1	1	1	1	1	1	0	1	83
17	0	0.658813	5691	1	1	1	1	1	1	1	1	0	0	0	1	66
18	1	0.391514	5773	0	1	1	1	1	0	0	0	0	0	0	1	39
19	0	0.030504	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:

   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 be 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	0.692600	2240	1	1	1	1	1	1	1	1	0	0	0	1	69
1	0	0.539248	113	1	1	1	1	1	1	1	0	0	0	0	1	54
2	1	0.718698	4132	1	1	1	1	1	1	1	1	1	0	0	1	72
3	0	0.133628	5573	0	1	1	0	0	0	0	0	0	0	0	0	13
4	0	0.014225	1109	0	1	0	0	0	0	0	0	0	0	0	0	1
5	0	0.031692	2282	0	1	0	0	0	0	0	0	0	0	0	0	3
6	1	0.854208	2976	1	1	1	1	1	1	1	1	1	1	0	1	85
7	0	0.399257	8431	0	1	1	1	1	0	0	0	0	0	0	1	40
8	1	0.745493	2770	1	1	1	1	1	1	1	1	1	0	0	1	75
9	1	0.995570	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.957687	2943	1	1	1	1	1	1	1	1	1	1	1	1	96
11	0	0.243037	1196	0	1	1	1	0	0	0	0	0	0	0	0	24
12	1	0.531214	8874	1	1	1	1	1	1	1	0	0	0	0	1	53
13	0	0.130486	1491	0	1	1	0	0	0	0	0	0	0	0	0	13
14	0	0.098482	7676	0	1	0	0	0	0	0	0	0	0	0	0	10
15	1	0.460246	8750	0	1	1	1	1	1	0	0	0	0	0	1	46
16	1	0.833093	5049	1	1	1	1	1	1	1	1	1	1	0	1	83
17	0	0.658813	5691	1	1	1	1	1	1	1	1	0	0	0	1	66
18	1	0.391514	5773	0	1	1	1	1	0	0	0	0	0	0	1	39
19	0	0.030504	3906	0	1	0	0	0	0	0	0	0	0	0	0	3