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 screeshot, the following are the top three variables that contribute most towards the probability of a lead getting converted :
   - a) Last Activity_SMS Sent
   - b) Total Time Spent on Website
   - c) Lead Source_Olark Chat

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.9651	0.020	47.685	0.000	0.925	1.005
TotalVisits	0.0243	0.006	4.344	0.000	0.013	0.035
Total Time Spent on Website	0.1984	0.005	36.733	0.000	0.188	0.209
Page Views Per Visit	-0.0251	0.007	-3.743	0.000	-0.038	-0.012
Lead Origin_API	-0.5913	0.022	-26.423	0.000	-0.635	-0.547
Lead Origin_Landing Page Submission	-0.6426	0.023	-28.413	0.000	-0.687	-0.598
Lead Origin_Lead Import	-0.6260	0.070	-8.895	0.000	-0.764	-0.488
Lead Source_Direct Traffic	-0.0537	0.013	-4.144	0.000	-0.079	-0.028
Lead Source_Olark Chat	0.1390	0.019	7.478	0.000	0.103	0.175
Last Activity_Converted to Lead	-0.1471	0.024	-6.095	0.000	-0.194	-0.100
Last Activity_Email Bounced	-0.2020	0.026	-7.735	0.000	-0.253	-0.151
Last Activity_Olark Chat Conversation	-0.2171	0.018	-11.820	0.000	-0.253	-0.181
Last Activity_Other	-0.0446	0.021	-2.166	0.030	-0.085	-0.004
Last Activity_Page Visited on Website	-0.1124	0.020	-5.568	0.000	-0.152	-0.073
Last Activity_SMS Sent	0.2078	0.012	17.594	0.000	0.185	0.231
Specialization_Finance Management	-0.0767	0.012	-6.211	0.000	-0.101	-0.053

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 :
   - 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 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 probabilty (value = 1) under a cut off 0.1 (column 0.1 highlighted).

	Converted	Converted_Prob	Lead Number	Converted_pred	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted
0	0	0.176616	609431	0	1	1	0	0	0	0	0	0	0	0	0
1	0	0.053991	631817	0	1	0	0	0	0	0	0	0	0	0	0
2	0	0.016089	596164	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0.181098	646570	0	1	1	0	0	0	0	0	0	0	0	0
4	0	0.028289	643974	0	1	0	0	0	0	0	0	0	0	0	0

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 probabilty (value = 1 highlighted in yellow color) under column 0.6. 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. This

should not be a major cause for concern as the target has already be achieved.

	Converted	Converted_Prob	Lead Number	Converted_pred	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	0.176616	609431	0	1	1	0	0	0	0	0	0	0	0	0	18
1	0	0.053991	631817	0	1	0	0	0	0	0	0	0	0	0	0	5
2	0	0.016089	596164	0	1	0	0	0	0	0	0	0	0	0	0	2
3	0	0.181098	646570	0	1	1	0	0	0	0	0	0	0	0	0	18
4	0	0.028289	643974	0	1	0	0	0	0	0	0	0	0	0	0	3
5	0	0.027199	614268	0	1	0	0	0	0	0	0	0	0	0	0	3
6	0	0.240730	633399	0	1	1	1	0	0	0	0	0	0	0	0	24
7	1	0.513379	615080	1	1	1	1	1	1	1	0	0	0	0	1	51
8	0	0.053991	627916	0	1	0	0	0	0	0	0	0	0	0	0	5
9	1	0.470405	622276	0	1	1	1	1	1	0	0	0	0	0	1	47
10	0	0.053991	646808	0	1	0	0	0	0	0	0	0	0	0	0	5
11	0	0.271117	597718	0	1	1	1	0	0	0	0	0	0	0	0	27
12	0	0.250522	613099	0	1	1	1	0	0	0	0	0	0	0	0	25
13	1	0.424125	607465	0	1	1	1	1	1	0	0	0	0	0	1	42
14	0	0.271117	642081	0	1	1	1	0	0	0	0	0	0	0	0	27
15	0	0.692123	613697	1	1	1	1	1	1	1	1	0	0	0	1	69
16	0	0.448216	638616	0	1	1	1	1	1	0	0	0	0	0	1	45
17	0	0.602906	615753	1	1	1	1	1	1	1	1	0	0	0	1	60
18	1	0.713385	615645	1	1	1	1	1	1	1	1	1	0	0	1	71
19	0	0.016590	589246	0	1	0	0	0	0	0	0	0	0	0	0	2