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

Ans:-

The below table shows all the features which are considered in our final model

	coef	std err	z	P>|z|	[0.025	0.975]
const	-2.2777	0.118	-19.289	0.000	-2.509	-2.046
Do Not Email	-1.2321	0.175	-7.050	0.000	-1.575	-0.890
Total Time Spent on Website	1.1164	0.040	27.570	0.000	1.037	1.196
Lead Origin_Lead Add Form	3.6582	0.223	16.409	0.000	3.221	4.095
Lead Origin_Lead Import	0.9391	0.450	2.085	0.037	0.056	1.822
Lead Source_Olark Chat	1.2609	0.104	12.107	0.000	1.057	1.465
Lead Source_Welingak Website	1.8629	0.760	2.451	0.014	0.373	3.353
Last Activity_Email Link Clicked	0.7517	0.240	3.137	0.002	0.282	1.221
Last Activity_Email Opened	1.3077	0.122	10.739	0.000	1.069	1.546
Last Activity_Other	1.8416	0.237	7.773	0.000	1.377	2.306
Last Activity_Page Visited on Website	0.8408	0.178	4.724	0.000	0.492	1.190
Last Activity_SMS Sent	2.4052	0.124	19.332	0.000	2.161	2.649
What is your current occupation_Other	-1.2082	0.087	-13.889	0.000	-1.379	-1.038
What is your current occupation_Working Professional	2.5505	0.188	13.563	0.000	2.182	2.919

We can see that the variables Lead Origin_Lead Add Form, What is your current occupation_Working Professional, Last Activity_SMS Sent are the top three variables which contributes towards the lead conversion probability.

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?

Ans:-

As seen in the above table the top three features obtained are the categorical/dummy variables. Hence, we can say that Lead Origin_Lead Add Form, What is your current occupation_Working Professional, Last Activity_SMS Sent are the top three categorical variables contributing towards the lead conversion.

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.

Ans:-

Based on the prediction made by the final model we have considered 0.35 probability as our optimal cutoff point to determine the predicted leads conversion. Since, the company wants to be more aggressive and convert almost all potential leads we can consider our probability as 0.1 by doing this we get leads which are predicted as having lower chance of conversion by our model.

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.

Ans:-

As discussed above we have considered 0.35 probability as our optimal cutoff point. We got to this point after plotting the accuracy, sensitivity and specificity at different probabilities. Since, the company has already reached it's quarterly goals and wants to save money by minimizing the phone calls we can consider 0.8 or 0.9 as our cutoff probability by doing this we get leads which have highly potential conversion rate as per our model.