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

   Ans> Below are the variables in the final model:

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.3170	0.057	-5.524	0.000	-0.429	-0.204
Do Not Email	-1.4293	0.162	-8.803	0.000	-1.748	-1.111
Total Time Spent on Website	0.9829	0.035	28.398	0.000	0.915	1.051
Landing Page Submission	-0.3908	0.071	-5.530	0.000	-0.529	-0.252
Lead Add Form	3.7831	0.218	17.367	0.000	3.356	4.210
Unknown Employment	-1.2578	0.082	-15.346	0.000	-1.418	-1.097
Working Professional	2.4713	0.179	13.776	0.000	2.120	2.823

   As we can see from above, Lead Origin (Lead Add Form), What is your current occupation (Working Professional, Unknown Employment) and Do Not Email are top three variables which contribute towards the probability of a lead getting converted as they have the highest absolute value of coefficients in the model.

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 per the final model shown above, below 3 categorical/ dummy variables should be focused upon to increase the probability of lead conversion:

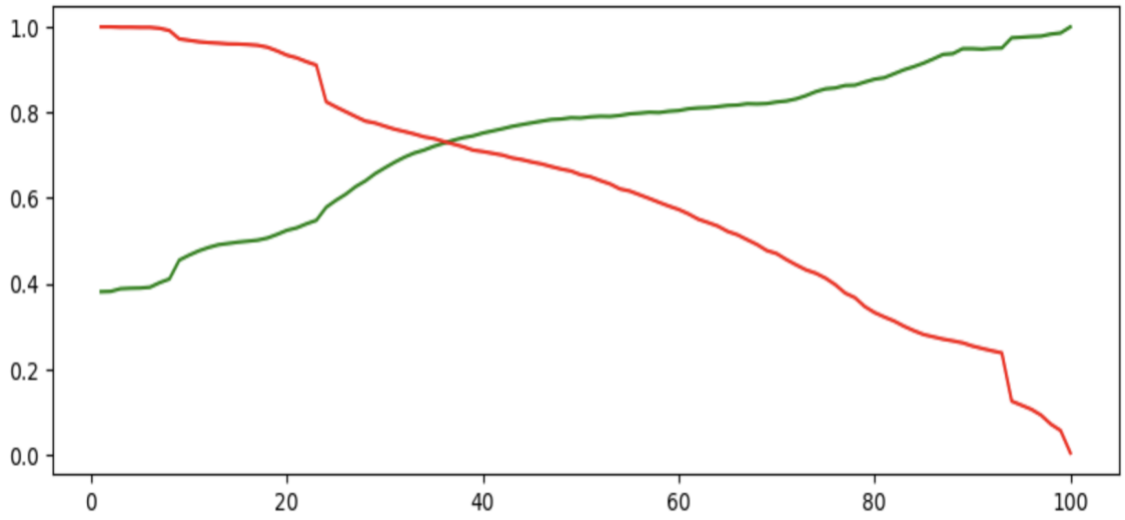
   - Lead Add Form- Leads originated via Lead Add form have the highest probability of conversion and should be focused upon.
   - Working Professional- Leads with employment status as Working Professional have the second highest probability of getting converted
   - Do Not Email- Leads which have selected Do No Email would most likely not avail the course and hence should not be focused upon (as negative coefficient in the model).

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> Since we are having additional resources in our team in form of interns, we suggest lowering the threshold of lead score from the existing value of 60 (for the required conversion rate of 80%). Since, as per the precision-recall curve, optimum cut off point is 40, we suggest taking the cut-off point as 40 to increase the number of probable leads we can call.

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> Given we need to reduce the number of useless phone calls, we need to maximize the conversion rate i.e. precision value for this case.



   As per the precision-recall curve shown above, we suggest taking the cut-off point for lead score as 80 which would result in precision value of 88% and hence result in only c. 12% incorrect calls.