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

   Ans: To determine the top three variables contributing most towards the probability of a lead getting converted, we should look at the coefficients which have highest positive value in the logistic regression model.
We see that 'Last activity', 'City' and 'Page Views Per Visit' contribute most as they have the highest positive coefficients.

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: 'City', 'Specialization' and 'Last notable activity' 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 because they have high coefficients (positive or negative) and low P value compared to others.

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: A good strategy that can be employed at this stage is
          a) To increase recall score by lowering the decision threshold for predicting conversions (e.g., 0.5 to 0.4), which will capture more leads and maximize the chances of contacting potential converts. By default, it uses a threshold of 0.5 to classify a lead as converted (1) or not converted (0). Lowering the decision threshold (e.g., from 0.5 to 0.3) will increase recall because the model will classify more leads as converted, capturing more true positives along with more false positives.

          b) We can segment the leads into different buckets named high confidence leads, medium confidence leads and low confidence leads.
Earlier, the CEO had given a ballpark of the target lead conversion rate to be around 80%.
So leads with lead score more than 80% should be put in high confidence leads segment, and focus should be heavy on this group with personal calls, emails, and follow-ups.
Similarly, the leads with lead score more than 60% should be put in medium confidence leads segment. These leads should be nurtured with slightly less aggressive follow-up, such as emails, automated responses, or informational videos.
Leads with lead score less than 60% can be put in low confidence leads segment. These leads can be deprioritized unless resources allow for further nurturing or follow-up.

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: In this case the precision score should be increased by raising the decision threshold. Raising the threshold (e.g., from 0.5 to 0.7) will increase precision because the model will classify leads as converted only when the probability is higher, reducing the number of false positives, along with reducing true positives. Thus the company can minimize the rate of useless phone calls.