1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
Ans. The top 3 variables that predict whether a lead will convert are as follows:

1. **Lead Origin_Lead Add Form** - when 'Lead Origin_Lead Add Form' takes the value 1 (compared to the reference 0), log odds of conversion increases by 2.856.
2. **What is your current occupation_Working Professional** - when 'What is your current occupation_Working Professional' takes the value 1 (compared to the reference 0), log odds of conversion increases by 2.4386.
3. **Lead Origin_Other** - when 'Lead Origin_Other' takes the value 1 (compared to the reference 0), log odds of conversion decreases by 1.8062

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. The top 3 categorical/dummy variables in the model which should be focused on the most to increase the probability of lead conversion are:

1. **Lead Origin_Lead Add Form** - leads which **Lead Origin** value equal to **Lead Add Form** should be prioritised.
2. **What is your current occupation_Working Professional** - **Working Professionals** should be prioritised because they are quite likely to convert.
3. **Country_Unknown** - leads whose **Country** is **Unknown** should be prioritised because they have higher conversion rates

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. Normally, the probability threshold of the logistic regression model sould be **0.4**. During the 2 months when interns are available to make more phone calls, the probability **threshold should be lowered** to classify more leads as ‘hot’.

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. Normally, the probability threshold of the logistic regression model sould be **0.4**. When the sales target is reached before the deadline and the company’s aim is to avoid unnecessary phone calls, the probability **threshold should be raised** to classify fewer leads as ‘hot’.