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

   Ans 1: The top three variables in model which contribute most towards the probability of lead getting converted are:
   a.Total Time Spent on Website
   b. Lead Origin
   c. What is your current occupation.

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 2: The top 3 categorical variables in the model which should be focused the most on in order to increase the probability of lead conversion are:
   a. Lead Source.
   b. Current Occupation.
   c. Last notable Activity

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 3: If all the potential leads (i.e. the customers who have been predicted as 1 by the model) is to be called then the potential leads need to prioritized in such a way that leads with highest probability should be called first and then followed by 2nd highest and so on.

   So, to prioritize the potential leads we should take the support of log odd of the model. The equation for log odds is:

   ln[P/(1-P)] = -2.79 + 3.8924*(Total Time Spent on Website) + 3.2760*(Lead Origin_Lead Add Form) – 0.5035*(Lead Source_Direct Traffic) + 2.4205*(Lead Source_Welingak Website) - 1.5070*(Do Not Email_Yes) + 1.3212*(Last Activity_SMS Sent) + 1.5297*(Last Activity_SMS Sent) + 2.4078*(What is your current occupation_Other) + 1.2552*(What is your current occupation_Student) + 1.1513*(What is your current occupation_Unemployed) + 3.6658*(What is your current occupation_Working Professional) – 0.8341*(Last Notable Activity_Olark Chat Conversation) + 1.8430*(Last Notable Activity_Unreachable).

With the above formula the log odd of all the potential lead will be calculated and prioritized as higher the log odd, higher is the probability. So the log odd should be sorted in descending order to get the hot leads.

Below are the variables which contribute most towards the probability of a lead getting converted, and leads with below variables should be called:

a. **Total Time Spent on Website** –Higher the value, higher is the probability of lead converting successfully. It has highest coefficient viz 3.8924.
b. **What is your current occupation** – Working professional looking to upgrade the skills have higher probability of lead converting successfully. Its coefficient is 3.6658
c. **Lead Origin** – If "Lead origin" = "Lead Add Form", probability of lead converting successfully is high. Its coefficient is 3.2760

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.

   Ans4: To minimize the rate of useless phone calls leads should be excluded based on the below variables and comment provided and should not be called

   a. Variable: Last Notable Activity – if Last Notable Activity" = "Olark Chat Conversation", then the lead should not be called as probability of successful conversion is less.

   b. Variable: Lead Source – if "Lead Source" = "Direct Traffic", then the lead should not be called as probability of successful conversion is less.

   c. Variable: Do Not Email – if "Do Not Email "= "Yes", then the lead should not be called as probability of successful conversion is less.