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

The top 3 variables are

- Lead Quality_High in Relevance
- Lead Origin_Lead Add Form
- Lead Source_Welingak Website

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?

Top 3 Categories are

- Lead Quality
- Lead Origin
- Lead Source

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.

When X Education has more man-power, they can decrease the cut-off from 40 to something like 35 or even 30 so the pool of possible leads increases. The code to show all leads with a Lead Score value of more than 30 is **“data[data["Lead Score"]>30]”**

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.

When X Education wants to decrease the man-power, they can increase the cut-off from 40 to something like 50 so the pool of possible leads decreases but the remaining leads have very high probablity of conversion.
The code to show all leads with a “Lead Score” value of more than 50 is “**data[data["Lead Score"]>50]”**