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

   Solution:
   Based on the coefficient values from below screeshot, the following are the top three variables that contribute most towards the probability of a lead getting converted :
   a) Total Time Spent on Website
   b) Lead Add Form (from Lead Origin)
   c) Working Professional (from What is your current occupation)

const	-2.8270
Lead Number	0.6844
TotalVisits	1.3003
Total Time Spent on Website	4.5675
Last Activity_Olark Chat Conversation	-1.3281
Last Activity_Other	-0.3341
Last Activity_Page Visited on Website	-0.5929
Lead Source_Google	0.2545
Lead Source_Olark Chat	1.5001
Lead Origin_Landing Page Submission	-0.2263
Lead Origin_Lead Add Form	4.2412
What is your current occupation_Working Professional	2.6484
Last Notable Activity_Modified	-0.5179
Last Notable Activity_SMS Sent	1.4472

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?

   Solution:
   Based on the Coefficient Values from the below screen shot, following are the top three categorical/dummy variables that should be focused the most in order to increase the probability of lead conversion:
   a) Lead Add Form (from Lead Origin)
   b) Working Professional (from What is your current occupation)
   c) Olark Chat (from Lead Source)

```
const                                                -2.8270
Lead Number                                           0.6844
TotalVisits                                           1.3003
Total Time Spent on Website                           4.5675
Last Activity_Olark Chat Conversation                -1.3281
Last Activity_Other                                  -0.3341
Last Activity_Page Visited on Website                -0.5929
Lead Source_Google                                    0.2545
Lead Source_Olark Chat                                1.5001
Lead Origin_Landing Page Submission                  -0.2263
Lead Origin_Lead Add Form                             4.2412
What is your current occupation_Working Professional  2.6484
Last Notable Activity_Modified                       -0.5179
Last Notable Activity_SMS Sent                        1.4472
```

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.

Solution:

In the below screenshot, the final prediction is calculated based on an optimal cut off value i.e. 0.37.

In order to make the Sales Aggressive, the company may contact all the leads which have a conversion probabilty (value = 1) under a cut off of 0.3 (column 0.3 highlighted in green).

```
y_train_pred_final[y_train_pred_final["final_predicted"] == 1].head(10)
```

	Converted	Lead_Prob	Converted_Pred	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
4	1	0.477830	0	1	1	1	1	1	0	0	0	0	0	1	48
15	1	0.850309	1	1	1	1	1	1	1	1	1	1	0	1	85
21	0	0.935704	1	1	1	1	1	1	1	1	1	1	1	1	94
23	1	0.996752	1	1	1	1	1	1	1	1	1	1	1	1	100
25	1	0.975761	1	1	1	1	1	1	1	1	1	1	1	1	98
26	1	0.990749	1	1	1	1	1	1	1	1	1	1	1	1	99
28	0	0.438951	0	1	1	1	1	1	0	0	0	0	0	1	44
29	0	0.417533	0	1	1	1	1	1	0	0	0	0	0	1	42
32	0	0.578936	1	1	1	1	1	1	1	0	0	0	0	1	58
36	1	0.973870	1	1	1	1	1	1	1	1	1	1	1	1	97

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.

Solution:

In order to Minimize the Rate of useless phone calls, the company can contact all the leads which have a Conversion Probabilty (value = 1 highlighted in green color) under column 0.8. However, we may include & miss out some of those leads that were not converted & are actually converted but the model wrongly predicted them as Converted & not converted. (See red highlights in the image below). This should not be a major drawback for this model as the target has already been achieved of 80% Conversion Rate.

```
y_train_pred_final[y_train_pred_final[0.7] == 1].head(10)
```

	Converted	Lead_Prob	Converted_Pred	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
15	1	0.850309	1	1	1	1	1	1	1	1	1	1	0	1	85
21	0	0.935704	1	1	1	1	1	1	1	1	1	1	1	1	94
23	1	0.996752	1	1	1	1	1	1	1	1	1	1	1	1	100
25	1	0.975761	1	1	1	1	1	1	1	1	1	1	1	1	98
26	1	0.990749	1	1	1	1	1	1	1	1	1	1	1	1	99
36	1	0.973870	1	1	1	1	1	1	1	1	1	1	1	1	97
45	1	0.866324	1	1	1	1	1	1	1	1	1	1	0	1	87
46	0	0.975318	1	1	1	1	1	1	1	1	1	1	1	1	98
49	0	0.761017	1	1	1	1	1	1	1	1	1	0	0	1	76
55	1	0.962976	1	1	1	1	1	1	1	1	1	1	1	1	96