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

Ans 1. The below image depicts the important features In the descending order: -

```
res.params.sort_values(ascending=False)

TotalVisits                                             9.866742
Total Time Spent on Website                             4.604821
Lead Origin_Lead Add Form                               3.750105
What is your current occupation_Working Professional    3.679731
Lead Source_Welingak Website                            2.582057
What is your current occupation_Other                   2.156716
Last Notable Activity_Unreachable                       1.815310
Lead Source_Olark Chat                                  1.580159
Last Activity_SMS Sent                                  1.267234
What is your current occupation_Student                 1.245642
What is your current occupation_Unemployed              1.163151
Last Activity_Olark Chat Conversation                  -1.397360
Do Not Email_Yes                                       -1.436043
const                                                  -3.453287
dtype: float64
```

So, the top 3 variables are: -
  i.   **Total Visits**
  ii.   **Total Time Spent on website**
  iii. **Lead Origin_Lead Add Form**

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. As can be seen from the image in previous question the dummy variables that should be focused the most on are: -
  a. **Lead Origin_Lead Add Form**
  b. **What is your current occupation_Working Professional**
  c. **Lead Source_Welingak Website**

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. A simple two step formula can be employed. The model will assign a **Lead_Score** to all of the Leads.
   a. It was observed that the **Lead_Score** greater than **36** had and **80%** chance to be converted to customer. Hence all of those leads should be aggressively contacted.
   b. The sales team is now divided into experienced and interns. The interns should be directed towards leads having **Lead_Score** on the higher end to secure the conversion whereas the one on the lower end are to be handled by the experienced.

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 4. A few steps for company would be to: -
   a. The company should work on improving their presence on different **Lead Sources** such as **Welingak_Website**, **Olark_Chat** etc.
   b. Company should work on **improving engagement on personal website** as these are major predictors for **Lead conversion**. Also, the form needs to be checked whether it can be improved and made further easy for the **Leads**.
   c. Company should also see if there are any changes to the parameters and update model accordingly. Not just that, but also see if the model needs an improvement in performance.