1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
   **Answer**:
   As per the coef values of the features in the final model shared below, here are the top 3 variables –
   1. Total Time Spent on Website
   2. Lead Source
   3. Lead Origin

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | -1.2226 | 0.134 | -9.144 | 0.000 | -1.485 | -0.961 |
| Do Not Email | -1.3560 | 0.158 | -8.599 | 0.000 | -1.665 | -1.047 |
| TotalVisits | 1.1889 | 0.349 | 3.409 | 0.001 | 0.505 | 1.872 |
| Total Time Spent on Website | 4.4597 | 0.155 | 28.681 | 0.000 | 4.155 | 4.764 |
| Lead Origin_Landing Page Submission | -0.9128 | 0.121 | -7.569 | 0.000 | -1.149 | -0.676 |
| Lead Origin_Lead Add Form | 3.5753 | 0.204 | 17.547 | 0.000 | 3.176 | 3.975 |
| Lead Source_Olark Chat | 0.8670 | 0.117 | 7.429 | 0.000 | 0.638 | 1.096 |
| Lead Source_Welingak Website | 2.5908 | 0.744 | 3.482 | 0.000 | 1.133 | 4.049 |
| Specialization_Hospitality Management | -1.0042 | 0.314 | -3.194 | 0.001 | -1.620 | -0.388 |
| Specialization_Others | -1.1240 | 0.116 | -9.666 | 0.000 | -1.352 | -0.896 |
| What is your current occupation_Working Professional | 2.6733 | 0.184 | 14.502 | 0.000 | 2.312 | 3.035 |

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?
   **Answer**:
   Referring to the coef values of the final model, here are the top 3 categorical/dummy variables –
   1. Lead Origin_Lead Add Form
   2. What is your current occupation_Working Professional
   3. 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.
   **Answer**:
   The model we've built not only categorizes the leads on the predicted conversion value of 0 and 1 but also assigns them a 'Lead Score', wherein higher the score means higher the probability of the conversion. So, to maximize the chance of lead conversion for the interns, I'd follow the below strategy/steps –
   a. First thing would be to run the model on the provided lead dataset and filter all the leads with the final_prediction value of 1.

b. Next, we can perform a random distribution of the filtered list across the 10 interns so all of them have a randomly equal chance of getting leads with high ‘Lead Score’
c. Then we ask each of the interns to sort the provided list to them in the descending order of the ‘Lead Score’ and contact the leads in the same order to maximize their chance of getting a conversion.

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.
**Answer**:
Given that the sales targets are already met, the interns can be leveraged in the below activities –
   a. Push for more sales by only concentrating on the top 10% or 5% of the leads based on the highest ‘Lead Score’ that way minimizing the effort on too many phone calls.
   b. Based on the final features of the model, identify the various aspects of lead source, origin, motivation factors and other data gathered to know where we should be investing more like targeting the Welingak website or making the X-Education’s website more informative and relevant content filled so that users spend more time on it as that is positively correlated to ‘Lead Score’.
   c. Work on email or other marketing campaigns to reach out to leads with lower ‘Lead Score’ to get their attention and in turn possibly increasing their ‘Lead Score’ in future.