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

The coefficients in a logistic regression model represent the influence of each variable on the log-odds of the dependent variable (in this case, lead conversion). Positive coefficients increase the likelihood of conversion, while negative coefficients decrease it.
Here are the **top three variables** based on the highest positive coefficients, meaning they contribute the most to increasing the probability of lead conversion:

1. **Lead Source_Welingak Website**:
    - Leads coming from the Welingak website are the most likely to convert.
2. **Lead Source_Reference**:
    - Leads that come through a reference source are also highly likely to convert.
3. **Total Time Spent on Website**:
    - The more time a lead spends on the website, the higher the likelihood of conversion.

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?**

Categorical or dummy variables represent specific categories of features, and focusing on these can help improve the conversion rate. The top three dummy variables with the most positive impact are:

1. **Lead Source_Welingak Website**:
    - Focusing on leads coming from the Welingak website can significantly boost conversions.
2. **Lead Source_Reference**:
    - Leads that are referred (e.g., word of mouth, partner referrals) also convert well. Increasing referral strategies could help.
3. **Last Notable Activity_Unreachable**:
    - Leads marked as unreachable may still be worth pursuing because, despite the challenges, they have a relatively high conversion rate.

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.**

During the 2-month period when X Education hires interns and wants to aggressively convert leads predicted as "1" (i.e., leads most likely to convert), here's an effective strategy:

- **Prioritize Phone Conversations**: According to the model, the variables:
  - **Last Notable Activity_Had a Phone Conversation**
  - **Last Activity_Had a Phone Conversation**

indicate that phone calls are associated with higher conversion rates. So, the interns should focus on leads that have already had or are due for phone conversations.

- **Target Welingak Website and Reference Leads**: Since leads from **Lead Source_Welingak Website** and **Lead Source_Reference** are the most likely to convert, interns should prioritize making calls to leads from these sources first.
- **Time Spent on Website**: Leads who have spent more time on the website (**Total Time Spent on Website**: should also be targeted, as they show strong intent and are likely to convert with personalized outreach.

In Summary:

1. Prioritize leads from the Welingak website and referrals.
2. Focus on making phone calls to those leads, especially those who have already shown engagement (e.g., high time spent on the website or past phone conversations).

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 the company reaches its sales target for the quarter and wants to reduce unnecessary calls, they should minimize calls to leads with low conversion likelihoods based on the following insights:

- **Avoid Calling Leads Marked as Unreachable**: The variable **Last Activity_Unreachable** and **Last Notable Activity_Unreachable** indicate that while these leads have some potential, the "Unreachable" tag might indicate difficulty in contacting. If the goal is to avoid wasting time, deprioritize these leads.
- **Do Not Focus on Certain Email-Based Activities**: Leads marked with activities like:
  - **Last Activity_Email Bounced**
  - **Last Notable Activity_Email Bounced**

  are less likely to convert, so deprioritizing them for phone calls can save effort.
- **Leads with 'Do Not Email'**: The variable **Do Not Email_Yes** (Coefficient = -1.134423) indicates these leads are less likely to engage, so avoid contacting them unless necessary.
- **Avoid Leads with Low Website Interaction**: Leads with low engagement on the website, especially those with low **Total Time Spent on Website** (Coefficient = 3.314191), are less likely to convert. Focus on leads who have spent significant time on the website.

In Summary:

1. Avoid leads marked as "Unreachable" or those with bounced emails.
2. Do not call leads who have opted out of email communication (Do Not Email).
3. Focus on leads that have shown strong online engagement and phone conversation activity.

By filtering out leads with lower coefficients (negative or close to zero), the sales team can efficiently reduce the rate of unnecessary calls.