SUBJECTIVE QUESTIONS AND ANSWERS

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

The top three variables in the logistic regression model that contribute most toward the probability of a lead getting converted are:

1. **Total Time Spent on Website**
    - The longer a lead spends on the website, the higher the likelihood they will convert. The coefficient for this variable is significant (positive), meaning an increase in time spent on the website correlates with a higher probability of conversion.
2. **Total Visits**
    - More visits by a lead correlate with higher conversion probability. The coefficient for TotalVisits is also significant (positive), indicating that leads who visit the site more frequently are more likely to convert.
3. **Lead Origin - Lead Add Form**
    - Leads originating from the "Lead Add Form" are more likely to convert compared to other lead origins. This variable has a positive coefficient, meaning it has a strong impact on increasing the probability 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?**

To The top three categorical/dummy variables to focus on are:

1. **Lead Source - Welingak Website**
    - Leads from this source are highly likely to convert, as this variable has a significant positive coefficient in the model. Focusing marketing efforts on leads from this source will increase conversion rates.
2. **Lead Source - Olark Chat**
    - Leads from Olark Chat also show a strong positive relationship with conversion. Enhancing engagement through chat services could drive more conversions.
3. **Do Not Email - Yes**
    - Surprisingly, leads who have opted for "Do Not Email" but are still engaged through other channels (like phone or chat) may have a higher conversion probability. Ensuring that these leads are not contacted through email but are pursued through other engagement channels might be key.

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 intern phase, when the focus is on aggressively converting leads, the following strategy should be employed:

- **Target All Predicted Leads (1's):**
  Since the model predicts a high probability of conversion for leads classified as "1," the sales team should **prioritize these leads** for immediate action. The primary strategy should be to **make phone calls to all leads predicted as "converted" (1)** to ensure these potential leads are nurtured and converted.
- **Prioritize High-Impact Variables:**
  Focus on the top contributing variables such as **Total Time Spent on Website** and **Total Visits**. Leads who spend significant time on the website or make multiple visits are more likely to convert and should be prioritized.
- **Increased Contact Frequency:**
  Since the period is short (2 months), the sales team should **increase the number of phone calls** to leads identified as high-probability, as well as **opt for follow-up calls** to ensure higher engagement and 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.**

When the company reaches its target before the deadline and wants to minimize unnecessary phone calls, the following strategy should be implemented:

- **Focus on Predicted Non-Converted Leads (0's):**
  Instead of calling all leads, focus on **leads predicted as "0" (not converted)** by the model, as these leads are less likely to convert. Avoid contacting leads who are already predicted as low-probability, unless there is new information that can influence conversion.
- **Risk-Weighted Calls:**
  Use a **threshold adjustment** in the model to call only leads that are above a specific risk threshold. This could be done by **setting a cutoff probability** (e.g., leads with a conversion probability above 0.6), ensuring that only leads with a reasonable chance of converting are contacted.
- **Emphasize Low-Cost Engagement Channels:**
  Rather than focusing on phone calls, consider using **email or automated outreach methods** to minimize operational costs. Use a combination of lower-touch channels for leads with low likelihood but higher engagement potential.
- **Reduce Frequency of Calls:**
  Implement a **"need-based" calling strategy** by limiting the number of calls made to only those leads that have interacted significantly (e.g., clicked on emails, visited the website recently, or shown a change in behavior) even if their predicted conversion probability is low.