SUMMARY

The types of leads that can be considered as "hot leads" based on the model's predictions have been outlined using a lead scoring process where a logistic regression model is built to predict whether a lead will get converted or not.

1. Features with High Impact:

- Attributes with a significant impact on lead conversion and help in distinguishing between hot and cold leads are:

  - 'Total Time Spent on Website'

  - 'Lead Source_Welingak Website'

  - 'Last Activity_SMS Sent'

  - 'What is your current occupation_Working Professional'

  - 'Last Notable Activity_Modified'

2. Key Attributes of Hot Leads:

- Leads who spend more time on the website indicates their higher engagement and interest in the offerings and are more likely to be hot leads.

- Leads generated from the 'Welingak Website' as the lead source are more likely to convert which could imply the effectiveness of this source in attracting genuinely interested leads.

- Leads with the last activity being 'SMS Sent' are more likely to convert. This suggests that following up with leads through SMS communication is effective in driving conversions.

- Leads who are 'Working Professionals' have a higher likelihood of converting. This indicates that individuals with a specific professional background are more likely to become customers.

- Leads associated with the 'Modified' last notable activity have a higher conversion probability. This might indicate that personalized follow-ups contribute to higher conversions.

3. Model's Performance:

- The model achieves a reasonably good accuracy, sensitivity, and specificity, suggesting that it is capable of identifying hot leads to a certain extent.

- The accuracy score is around 80-81%, indicating that the model's predictions are accurate for a significant portion of the dataset.

- The sensitivity score is also around 80%, which means that the model is able to correctly identify a substantial proportion of actual converting leads.

- The specificity score is around 81%, implying that the model can effectively distinguish between converting and non-converting leads.

4. Optimal Cutoff Point:

- An optimal cutoff point of around 0.38 based on the ROC curve analysis is obtained which balances accuracy, sensitivity, and specificity to make predictions.

The "hot leads" can be identified by considering the specific features and attributes highlighted by the Machine Learning model, with Leads that exhibit higher engagement (e.g., spending more time on the website), coming from specific sources and having specific professional backgrounds are more likely to convert.