Hot Leads Identification Report

-Sindhu N

-Aniket Verma

Problem Statement

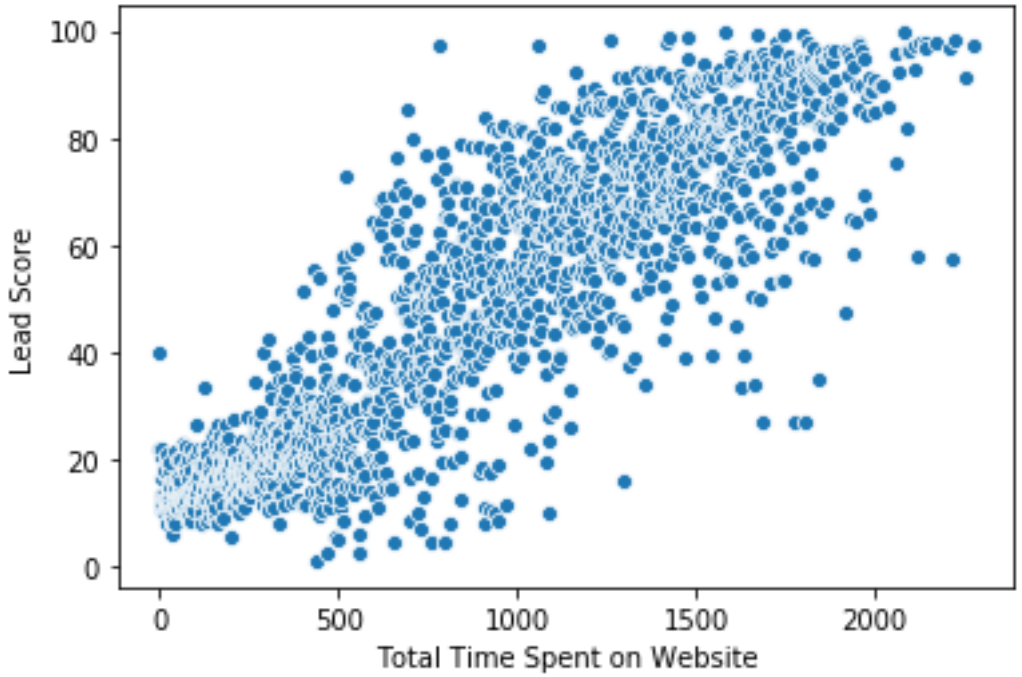
- An online education provider company ‘X Education’ wants to identify potential leads which they feel have a very strong possibility of being associated with any of the courses they offer.
- Our goal is to analyse the data provided by X Education and provide them the lead scores for each possible lead which would tell them how likely is the lead to take up the courses offered by them.

*A lead is a person who has expressed some sort of interest in the courses and has a certain probability of subscribing / purchasing the same.

Analysis Approach

- We use logistic regression to solve this problem.
- The aim of the logistic regression model is to determine the parameters which when focussed upon increases the probability of a lead to be converted.
- We generate lead scores for every lead in the data set. Higher the lead score, higher the probability of a lead to be converted.
- We then observe how the parameters determined by the model affect the lead score.

Total Time Spent On the Website

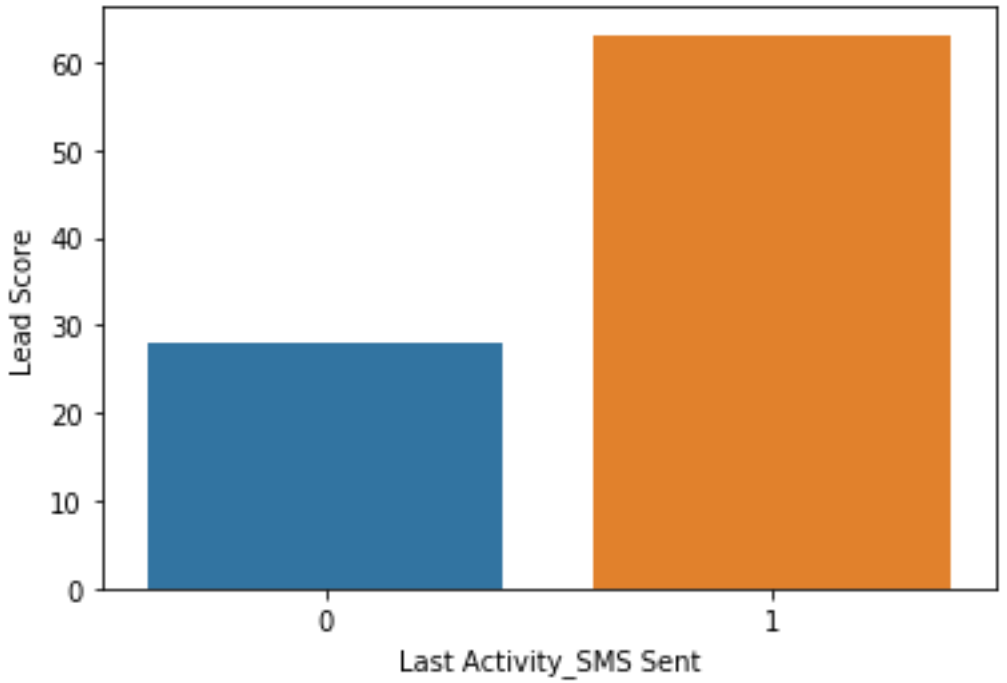


Total Time Spent On the Website

- Observation: We can observe that as the total time spent on the website increases, the lead score increases. The more a person spends time browsing the website of X Education, the more the chances are that he / she might opt for the programs offered by them.

- Inference: X Education needs to polish their website and provide as comprehensive guidance on their website as possible. The potential lead spends a lot of time browsing the website and hence get most of the information from it.

Last Activity (SMS Sent)

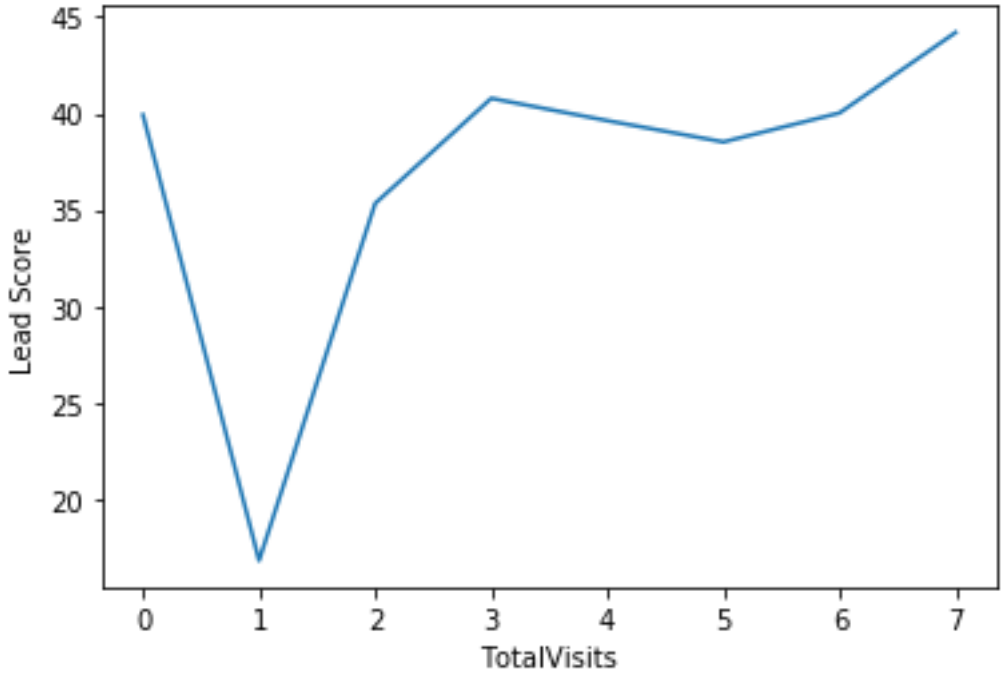


Last Activity (SMS Sent)

- Observation: We can observe that if a person has tried to contact the X Education team via an SMS, its more likely that he / she would opt for the services.

- Inference: People who contact via SMS should be closely and consistently followed up as they may prove to be potential converted leads.

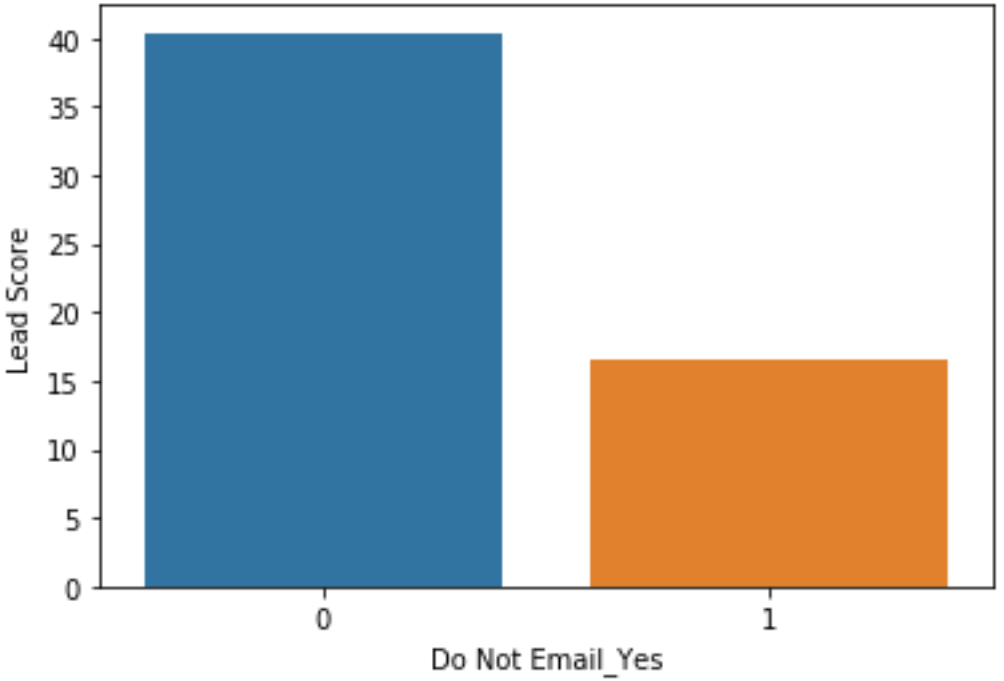
Total Visits on the Website



Total Visits on the Website

- Observation: We observe that the probability of a person who has done more visits on the website is more likely to be interested. However, certain anomalies like the decline at 1 visit from 0 can be observed. This can be attributed to the fact that a major chunk of the people would have landed on the website just out of curiosity and after finding that it wasn’t relevant for them, they never came back. Such leads can be safely ignored as the program doesn’t align with their career goals.
- Inference: Similar to the time spent on the website. The company could invest more resources in enhancing User Experience of the website along with the content.

Do Not Email

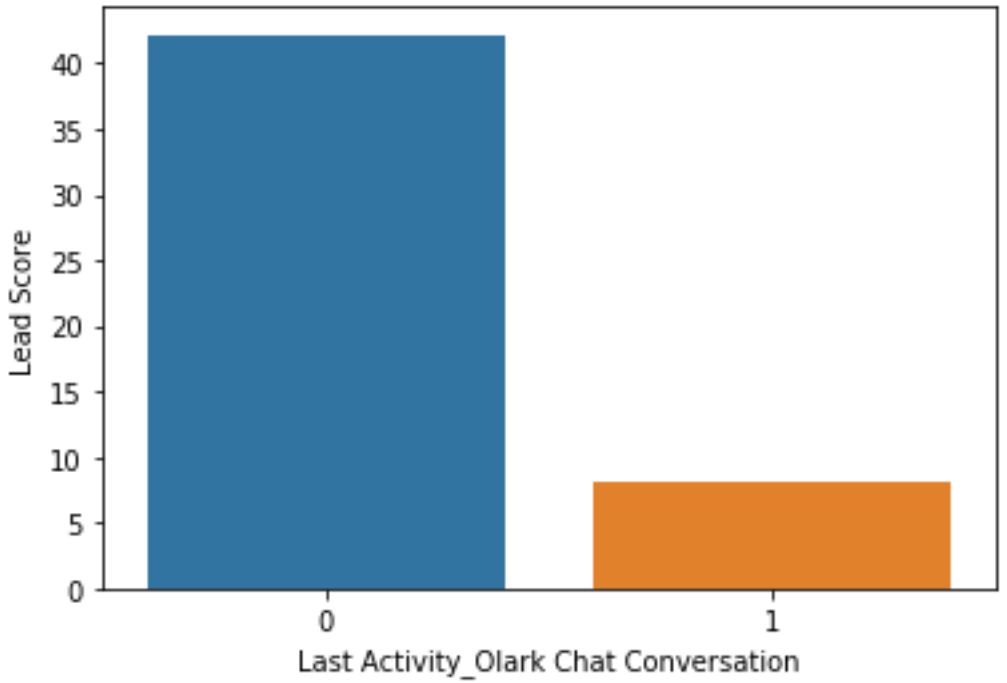


Do Not Email

- Observation: The trend here is quite straightforward. If a person subscribes to mailing alerts by the program, it indicates that the person is keen to find out more and is very likely to opt for the courses after brainstorming a little.

- Inference: To subscribe or not to subscribe is entirely a person's decision and cannot be directly influenced by the company. However, they can still roll out periodic discount packages and alert the user of any new additions to the program which may capture the person's interest.

Last Activity (Chat Conversation)



Last Activity (Chat Conversation)

- Observation: This is very interesting trend where people who have a chat conversation are way less likely to show interest. In fact this mode of communication is almost fatal to the organisation.

- Inference: Its hard to predict what could actually go wrong here. However, the possible scenarios could be
  - They might be talking to an AI chatbot which might not be capable enough to provide a dedicated feedback on the program.
  - They delegated person on the other end of the chat might not be equipped with sufficient information, thereby hampering user experience.

Thank You