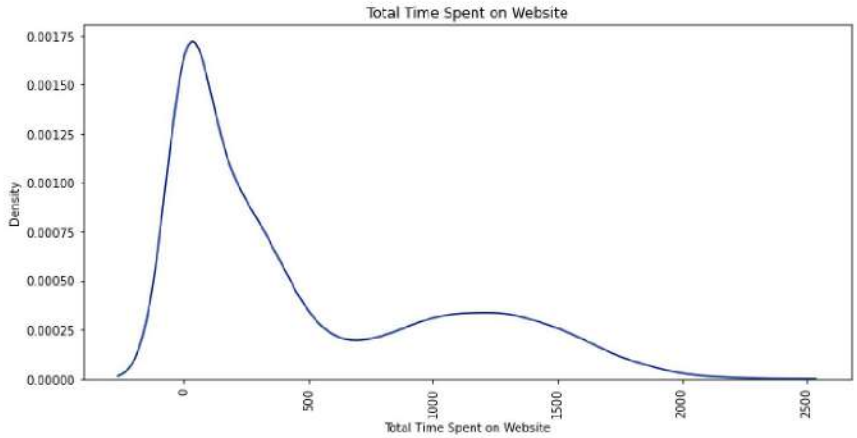
Subjective Questions and Answers

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

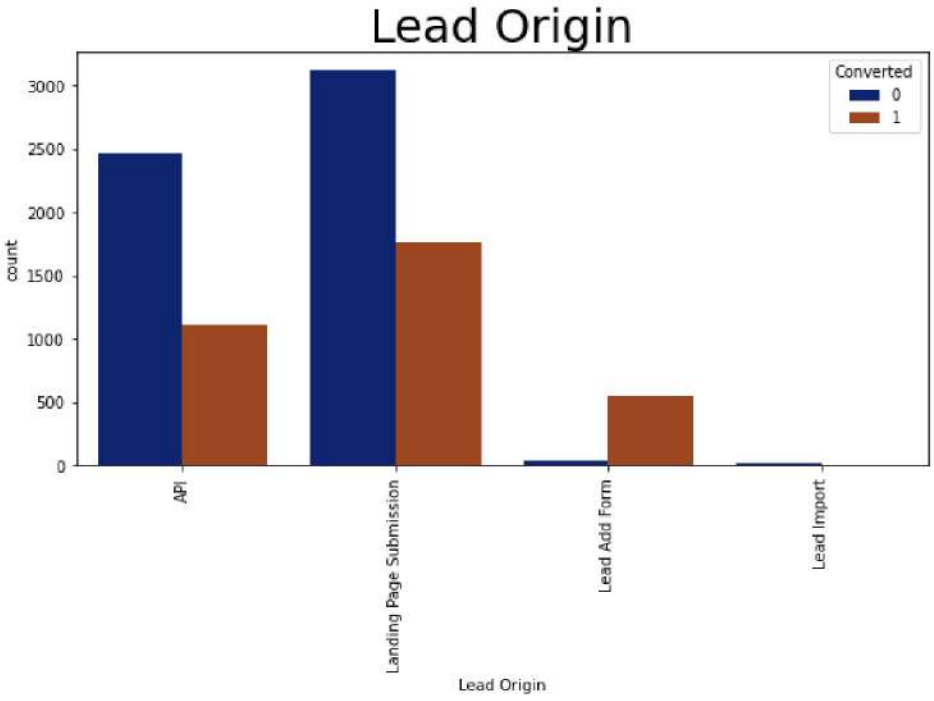
Solution:

The top three variables we found in the model, that contribute towards lead conversion are:

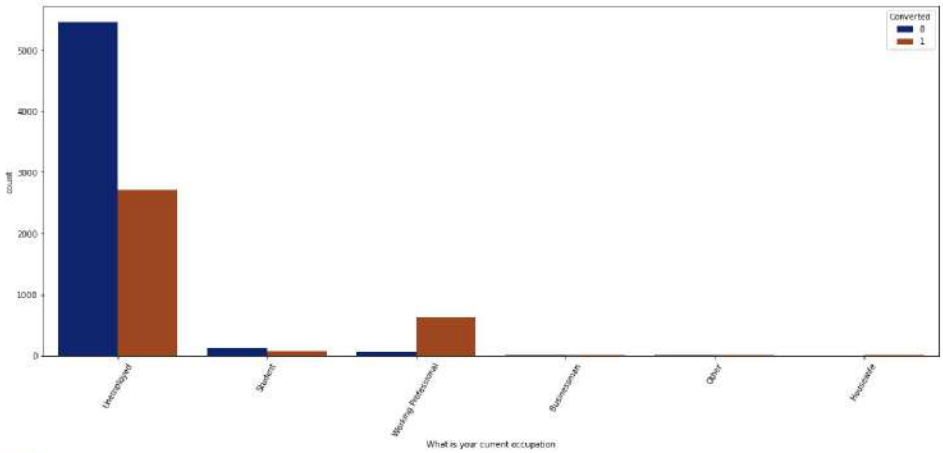
1. Total Time Spent on Website



2. Lead Origin_Lead Add Form



3. What is your current occupation_Working Professional



Conclusion: So, we can conclude that Total Time Spent on Website is the major contributing factor towards the 'Hot Leads'.

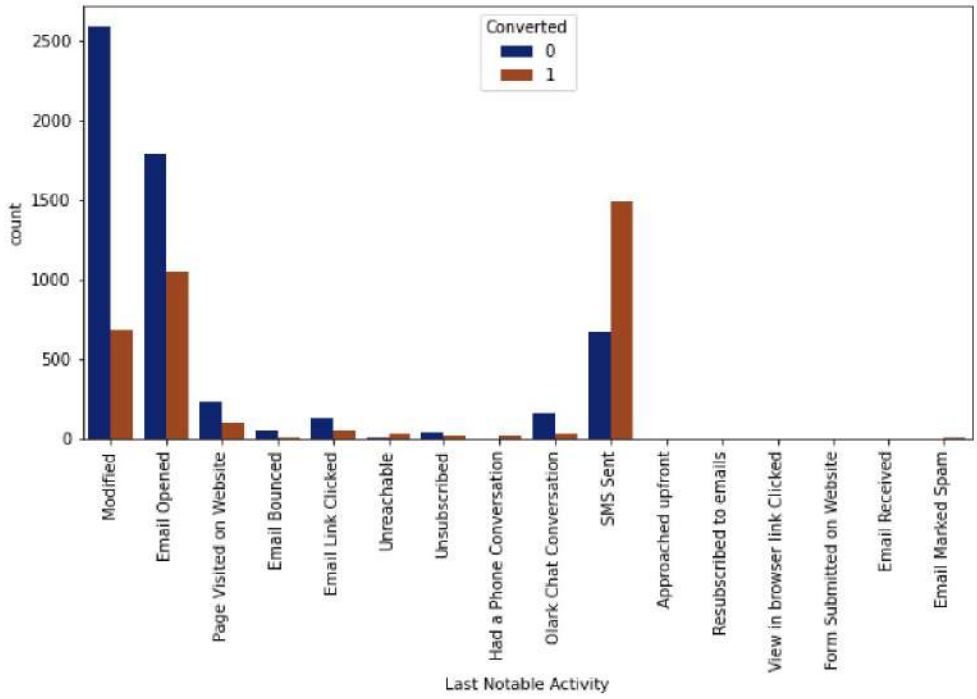
---

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?**

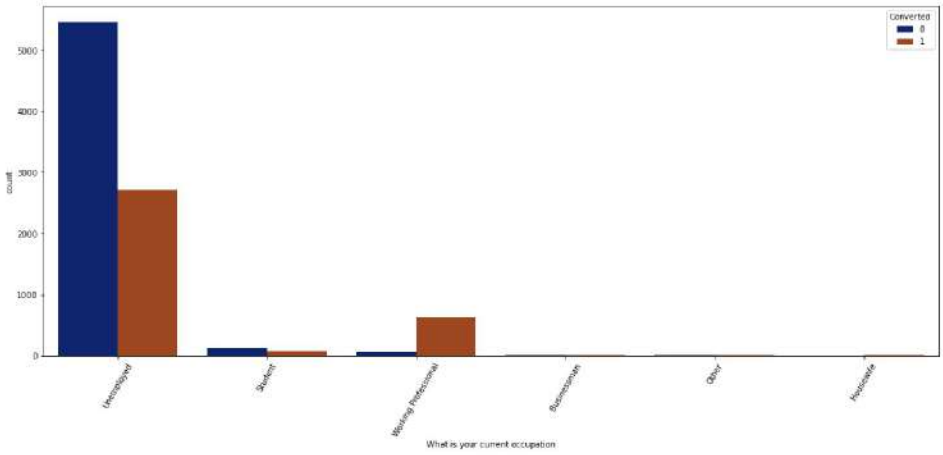
Solution:

The top three variables in my model, that should be focused are:

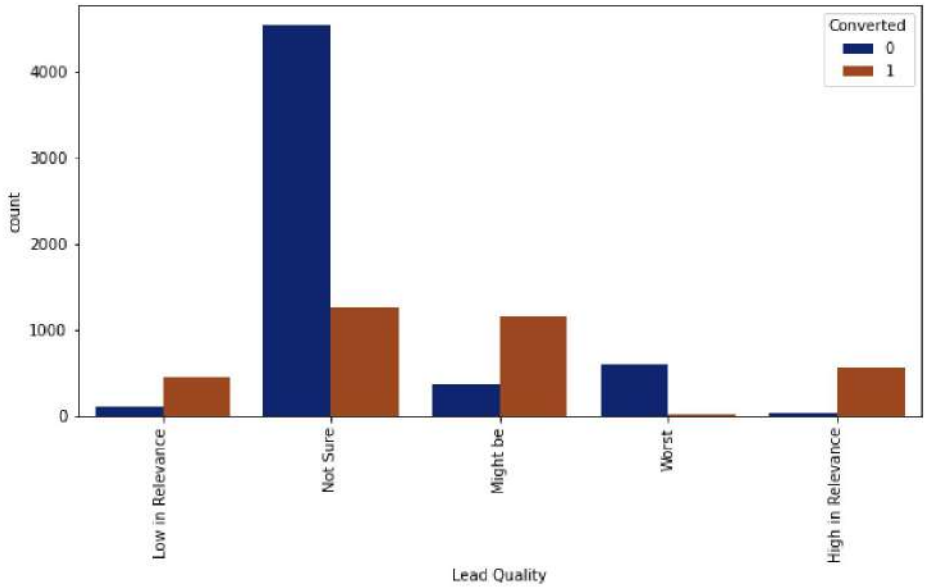
1. Last Notable Activity_SMS Sent (positively impacting)



2. What is your current occupation_Working Professional



3. Lead Quality_worst(negatively impacting)



	Features	VIF
16	What is your current occupation_Unemployed	33.42
25	Tags_Will revert after reading the email	12.52
28	Lead Quality_Not Sure	8.51
1	Lead Origin_Landing Page Submission	7.52
36	Last Notable Activity_SMS Sent	7.13
9	Last Activity_SMS Sent	6.82
13	Specialization_Others	5.03
17	What is your current occupation_Working Profes...	3.69
24	Tags_Ringing	3.57
33	Last Notable Activity_Modified	3.46
27	Lead Quality_Might be	2.85
4	Lead Source_Olark Chat	2.39
29	Lead Quality_Worst	2.28
2	Lead Origin_Lead Add Form	2.26
7	Last Activity_Olark Chat Conversation	2.15
21	Tags_Interested in other courses	1.94
19	Tags_Closed by Horizzon	1.93
15	What is your current occupation_Student	1.74
10	Last Activity_Unreachable	1.70
37	Last Notable Activity_Unreachable	1.65
6	Last Activity_Email Bounced	1.52
26	Tags_switched off	1.50
23	Tags_Other_Tags	1.49
0	Total Time Spent on Website	1.45
5	Lead Source_Welingak Website	1.42
18	Tags_Busy	1.41
34	Last Notable Activity_Olark Chat Conversation	1.40
22	Tags_Lost to EINS	1.36
31	Last Notable Activity_Email Bounced	1.34
20	Tags_Interested in full time MBA	1.19
35	Last Notable Activity_Page Visited on Website	1.13
32	Last Notable Activity_Email Link Clicked	1.07
14	Specialization_Travel and Tourism	1.05
12	Specialization_International Business	1.04
3	Lead Origin_Lead Import	1.04
11	Specialization_E-COMMERCE	1.03
30	City_Tier II Cities	1.03
8	Last Activity_Others	1.02

Conclusion: So, focus on sending more SMS notifications and improve the Lead Quality service in order to increase the probability of lead conversion.

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.**

Solution:

A good strategy will be:

- To focus on wider set of lead audience (inclusion of slightly lower conversion probable leads)
- Technically, we can generate this new set of leads by altering (moving down) the value of cut off so as to include more leads as the hot leads from our Logistic Regression Model
- Doing so, we will be better utilizing resourses and improving chance of converting a lead whose lead conversion probability might be low as well.

---

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.**

Solution:

A good strategy will be:

- To focus on narrow set of lead audience (discarding lower conversion probable leads)
- Technically, we can generate this new set of leads by altering (moving up) the value of cut off so as to discard lower conversion rate prabable leads from our Logistic Regression Model
- Doing so, we will be doing minimal effort and still be getting fair conversions.