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

Dep. Variable:	Converted	No. Observations:	3728
Model:	GLM	Df Residuals:	3714
Model Family:	Binomial	Df Model:	13
Link Function:	Logit	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-1741.0
Date:	Sun, 12 Mar 2023	Deviance:	3482.0
Time:	18:38:56	Pearson chi2:	3.95e+03
No. Iterations:	7	Pseudo R-squ. (CS):	0.3420
Covariance Type:	nonrobust		

	coef	std err	z	P> z	[0.025	0.975]
const	0.3648	0.269	1.356	0.175	-0.162	0.892
TotalVisits	0.4571	0.159	2.882	0.004	0.146	0.768
Total Time Spent on Website	4.0353	0.183	22.093	0.000	3.677	4.393
Lead Source_Reference	5.3182	0.601	8.852	0.000	4.141	6.496
Do Not Email_Yes	-1.4666	0.227	-6.448	0.000	-1.912	-1.021
Last Activity_Email Link Clicked	0.5891	0.300	1.965	0.049	0.002	1.177
Last Activity_Email Opened	0.7175	0.128	5.608	0.000	0.467	0.968
Last Activity_Had a Phone Conversation	2.7923	0.774	3.609	0.000	1.276	4.309
Last Activity_SMS Sent	1.5496	0.130	11.906	0.000	1.295	1.805
Last Activity_Unreachable	0.6283	0.385	1.634	0.102	-0.125	1.382
Last Activity_Unsubscribed	1.8260	0.525	3.477	0.001	0.797	2.855
Specialization_Hospitality Management	-0.6046	0.331	-1.824	0.068	-1.254	0.045
What is your current occupation_Student	-2.4377	0.372	-6.556	0.000	-3.166	-1.709
What is your current occupation_Unemployed	-3.2356	0.244	-13.253	0.000	-3.714	-2.757

Based on the coefficient values from above scree shot, the following are the top three variables that contribute most towards the probability of a lead getting converted:

- Reference(Lead Source)
- Total time spent on website
- Had a phone conversation(Lead Activity)

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?

Based on the coefficient values from the screen shot in the question above, the following are the top three categorical/dummy variables that should be focused in order to increase the probability of lead conversion

- Reference(Lead Source)
- Had a phone conversation(Lead Activity)
- Unsubscribed(Lead Activity)
- 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.

The below image has the final prediction which calculated based on a optimal cut off value of 0.4. In order to make the sales aggressive, the company may contact all the leads which have a conversion probability (value = 1) under a cut off 0.4 (column 0.4 highlighted in yellow).

	Converted	Prediction	Lead ID	Converted_Prediction	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	8.0	0.9	Final Prediction
0	0	0.096429	7813	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.168365	8549	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0.518277	2059	1	1	1	1	1	1	1	1	1	1	1	1
3	1	0.882857	8745	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0.225576	530	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0.961090	2473	1	1	1	1	1	1	1	1	1	1	1	1
6	0	0.018893	2993	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0.143110	3799	0	0	0	0	0	0	0	0	0	0	0	0
8	1	0.692107	188	1	1	1	1	1	1	1	1	1	1	1	1
9	0	0.277001	2401	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0.102011	4823	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0.178934	1485	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0.300280	966	0	0	0	0	0	0	0	0	0	0	0	0
13	1	0.991268	8243	1	1	1	1	1	1	1	1	1	1	1	1
14	1	0.710751	1655	1	1	1	1	1	1	1	1	1	1	1	1
15	1	0.140909	35	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0.164747	103	0	0	0	0	0	0	0	0	0	0	0	0
17	1	0.655689	4604	1	1	1	1	1	1	1	1	1	1	1	1
18	0	0.152654	3226	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0.699419	1250	1	1	1	1	1	1	1	1	1	1	1	1

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.

For the company to save time on its resources, the company may contact all the leads which have a conversion probability (value = 1) under column 0.7. However, the down side here is that they will miss out on those leads that are actually converted but then the prediction was wrong. But since the target is achieved it can be ignored

	Converted	Prediction	Lead ID	Converted_Prediction	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	8.0	0.9	Final Prediction
0	0	0.096429	7813	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.168365	8549	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0.518277	2059	1	1	1	1	1	1	1	1	1	1	1	1
3	1	0.882857	8745	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0.225576	530	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0.961090	2473	1	1	1	1	1	1	1	1	1	1	1	1
6	0	0.018893	2993	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0.143110	3799	0	0	0	0	0	0	0	0	0	0	0	0
8	1	0.692107	188	1	1	1	1	1	1	1	1	1	1	1	1
9	0	0.277001	2401	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0.102011	4823	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0.178934	1485	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0.300280	966	0	0	0	0	0	0	0	0	0	0	0	0
13	1	0.991268	8243	1	1	1	1	1	1	1	1	1	1	1	1
14	1	0.710751	1655	1	1	1	1	1	1	1	1	1	1	1	1
15	1	0.140909	35	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0.164747	103	0	0	0	0	0	0	0	0	0	0	0	0
17	1	0.655689	4604	1	1	1	1	1	1	1	1	1	1	1	1
18	0	0.152654	3226	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0.699419	1250	1	1	1	1	1	1	1	1	1	1	1	1