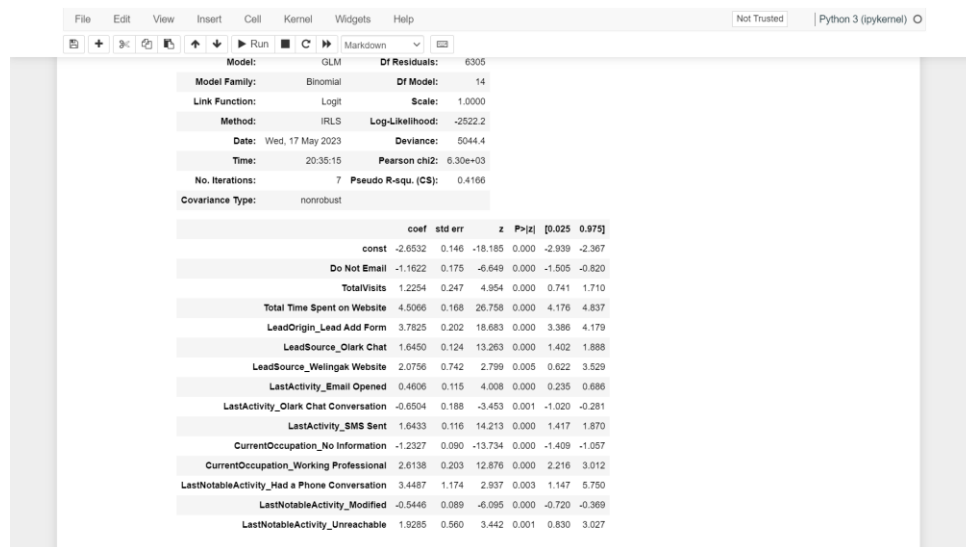


- Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

The Top 3 variables that contribute most towards the probability of a lead getting converted are

- Total Time Spent on Website**
- LeadOrigin\_Lead Add Form**
- LastNotableActivity\_Had a Phone Conversation**



	coef	std err	z	P> z	[0.025 0.975]
const	-2.6532	0.146	-18.185	0.000	-2.939 -2.367
Do Not Email	-1.1622	0.175	-6.649	0.000	-1.505 -0.820
TotalVisits	1.2254	0.247	4.954	0.000	0.741 1.710
Total Time Spent on Website	4.5066	0.168	26.758	0.000	4.176 4.837
LeadOrigin_Lead Add Form	3.7825	0.202	18.683	0.000	3.386 4.179
LeadSource_Olark Chat	1.6450	0.124	13.263	0.000	1.402 1.888
LeadSource_Welink Website	2.0756	0.742	2.799	0.005	0.622 3.529
LastActivity_Email Opened	0.4606	0.115	4.008	0.000	0.235 0.686
LastActivity_Olark Chat Conversation	-0.6504	0.188	-3.453	0.001	-1.020 -0.281
LastActivity_SMS Sent	1.6433	0.116	14.213	0.000	1.417 1.870
CurrentOccupation_No Information	-1.2327	0.090	-13.734	0.000	-1.409 -1.057
CurrentOccupation_Working Professional	2.6138	0.203	12.876	0.000	2.216 3.012
LastNotableActivity_Had a Phone Conversation	3.4487	1.174	2.937	0.003	1.147 5.750
LastNotableActivity_Modified	-0.5446	0.089	-6.095	0.000	-0.720 -0.369
LastNotableActivity_Unreachable	1.9285	0.560	3.442	0.001	0.830 3.027

- 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?

Again, 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 the most in order to increase the probability of lead conversion :

- Lead Add Form (from Lead Origin)
- Had a Phone Conversation ( from Last Notable Activity)
- Working Professional ( from What is your current occupation)

- 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.

```

In [123]: 1 # Now Let us calculate the Lead score
          2
          3 y_train_pred_final['lead_score'] = y_train_pred_final.Converted_Prob.map(lambda x: round(x*100))
          4 y_train_pred_final.head(20)

Out[123]:

```

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
0	0	0.701779	2240	1	1	1	1	1	1	1	1	1	0	0	1	70
1	0	0.591069	113	1	1	1	1	1	1	1	0	0	0	0	1	59
2	1	0.765682	4132	1	1	1	1	1	1	1	1	1	0	0	1	77
3	0	0.129061	5573	0	1	1	0	0	0	0	0	0	0	0	0	13
4	0	0.013213	1109	0	1	0	0	0	0	0	0	0	0	0	0	1
5	0	0.030126	2282	0	1	0	0	0	0	0	0	0	0	0	0	3
6	1	0.869418	2976	1	1	1	1	1	1	1	1	1	1	0	1	87
7	0	0.460064	8431	0	1	1	1	1	1	0	0	0	0	0	1	46
8	1	0.739865	2770	1	1	1	1	1	1	1	1	1	0	0	1	74
9	1	0.995442	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.958623	2943	1	1	1	1	1	1	1	1	1	1	1	1	96
11	0	0.263791	1196	0	1	1	1	0	0	0	0	0	0	0	0	26
12	1	0.522613	8874	1	1	1	1	1	1	1	0	0	0	0	1	52

Based on the above image where the final predictions were done based on the optimal cutoff point of 0.37.

Thus, the Sales team can contact the leads that are predicted with value 1 under 0.3 cutoff point.

- 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.

To minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1) based on the cutoff points of 0.6 or 0.7 based on the image below. However, we may miss out on those leads that are converted but then the model wrongly predicted them as not converted. This should not be a major cause for concern as the target has already been achieved.

```

In [123]: 1 # Now Let us calculate the Lead score
          2
          3 y_train_pred_final['lead_score'] = y_train_pred_final.Converted_Prob.map(lambda x: round(x*100))
          4 y_train_pred_final.head(20)

Out[123]:

```

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted	lead_score
0	0	0.701779	2240	1	1	1	1	1	1	1	1	1	0	0	1	70
1	0	0.591069	113	1	1	1	1	1	1	1	0	0	0	0	1	59
2	1	0.765682	4132	1	1	1	1	1	1	1	1	1	0	0	1	77
3	0	0.129061	5573	0	1	1	0	0	0	0	0	0	0	0	0	13
4	0	0.013213	1109	0	1	0	0	0	0	0	0	0	0	0	0	1
5	0	0.030126	2282	0	1	0	0	0	0	0	0	0	0	0	0	3
6	1	0.869418	2976	1	1	1	1	1	1	1	1	1	1	0	1	87
7	0	0.460064	8431	0	1	1	1	1	1	0	0	0	0	0	1	46
8	1	0.739865	2770	1	1	1	1	1	1	1	1	1	0	0	1	74
9	1	0.995442	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.958623	2943	1	1	1	1	1	1	1	1	1	1	1	1	96
11	0	0.263791	1196	0	1	1	1	0	0	0	0	0	0	0	0	26
12	1	0.522613	8874	1	1	1	1	1	1	1	0	0	0	0	1	52