ASSIGNMENT SUBJECTIVE QUESTIONS

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

ANS: First we need to have a glance on our model variables and its coefficient.

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.2254	0.130	-1.729	0.084	-0.481	0.030
Do Not Email	-1.4509	0.196	-7.397	0.000	-1.835	-1.067
Total Time Spent on Website	1.0481	0.041	25.580	0.000	0.968	1.128
Lead Origin_Landing Page Submission	-1.0627	0.130	-8.164	0.000	-1.318	-0.808
Lead Source_Olark Chat	0.9044	0.122	7.439	0.000	0.666	1.143
Lead Source_Reference	3.4792	0.246	14.122	0.000	2.996	3.962
Lead Source_Welingak Website	6.7192	1.018	6.600	0.000	4.724	8.715
Specialization_Hospitality Management	-0.8878	0.342	-2.594	0.009	-1.559	-0.217
Specialization_Unknown	-1.2125	0.125	-9.724	0.000	-1.457	-0.968
What is your current occupation_Student	0.2501	0.230	1.087	0.277	-0.201	0.701
What is your current occupation_Working Professional	2.5948	0.201	12.888	0.000	2.200	2.989
Last Notable Activity_Modified	-0.6043	0.086	-7.032	0.000	-0.773	-0.436
Last Notable Activity_Olark Chat Conversation	-1.1951	0.326	-3.664	0.000	-1.834	-0.556
Last Notable Activity_SMS Sent	1.4669	0.089	16.492	0.000	1.293	1.641
Last Notable Activity_Unreachable	1.8463	0.610	3.027	0.002	0.651	3.042
Last Notable Activity_Unsubscribed	1.1953	0.528	2.264	0.024	0.160	2.230

Based on the model we can infer that the top three variables are

1. Lead Origin
2. What is your current Occupation
3. Last notable activity

These three variables give us most of the inference and helps in calculating the lead score. By targeting these variables and analyzing the coefficient of it categorical dummy variables will increase the lead score which contribute most towards the probability of a lead getting converted. high positive coefficient values and this indicates higher the coefficient higher the possibility of increase the probability of lead conversion. If there is negative coefficient value, then it indicates that if these variables are targeted then the lead score will be decreased. So, we can try to reduce these values.

QUESTION 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?

ANS: Based on the results obtained by our model the top three categorical/dummy variables are

1. Lead origin – Welingak Website
2. Lead origin – Reference
3. What is your current occupation- Working

These three are chosen has it has high positive coefficient values and this indicates higher the coefficient higher the possibility of increase the probability of lead conversion

.

QUESTION 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.

ANS: The business can reduce the probability cut off from 0.35 i.e., can target customers with lead score less than 35 say 0.20 i.e. can target customers with lead score less than 20. Our model sensitivity was 80% when cut off was 0.35 and the sensitivity of the model is 88% when cut off is 0.2%. This change would lead to an increase in the targeted customers, and it might result in high lead conversion rate in turn improving the profits of the business.

QUESTION 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.

ANS: The business can choose to increase the cut off from 0.35 to a higher one say 0.5. The sensitivity of the model was 80% when cut off was 0.35 and the sensitivity of the model is 68% when cut off is 0.5%. This strategy can be applied when the business has reached its target and is willing to accept less lead conversion rate or if the business wants to concentrate on any other vertical and intends to reduce the sales team work (thereby diverting the employees towards a new vertical, if any. During the application of this strategy, make sure that you target at least those with lead source of welingak and reference categories and working professionals as these have very high prospects of getting hot leads and potential paying customers.