Lead Score Assignment

<u>Objective:</u>

An Education company X Education has appointed you to help them select the most promising leads. company requires you to build a model wherein you need to assign a lead score to each of the leadssuch that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.

Below mentioned are the steps to fulfill the objective of this assignment:

Reading and cleaning the data:

- Read the data and performed the basic checks e.g. info, shape , describe.
- To clean the dataset we removed the redundant data e.g. Prospect ID','Lead Number', 'Country', 'I agree to pay the amount through cheque' and 'a free copy of Mastering The Interview'.
- Dropped the columns having more than 30% of missing values.
- Few of the null values were changed to 'not provided' so as to not lose much data.

EDA:

- A quick EDA was done to check the condition of our data. It was found that a lot of elements in the categorical variables were irrelevant. The numeric values seem to be good.
- Changed the multicategory labels into dummy variables and binary variables into '0' and '1'.
- Checked the outliers and created bins for them as they were important variables and cannot be removed.
- Removed all the redundant and repeated columns from the data set for further analysis.

Data Preparation:

- Split the dataset into train and test dataset and scaled the dataset.
- Post this, we have plotted a heatmap to check the correlations among the variables.

Model Building:

- Created model with rfe count 19 and 15 and compared the model and choose our final model with rfe 19 variables as has more stability and accuracy than the other.
- Checked the precision and recall with accuracy, sensitivity and specificity for our final model and the tradeoffs.
- We found the score of accuracy and sensitivity from our final test model is in acceptable range. Accuracy 81%, Precision 74% and Recall 79%.

Conclusion:

- Important features responsible for good conversion rate or the ones' which contributes more towards the probability of a lead getting converted are :
  - Last Notable Activity_Had a Phone Conversation
  - What is your current occupation_Working Professional
  - Lead Origin_Lead Add Form

.