greatlearning

Problem Statement – Practice Project

General Instructions:

1. **A LMN medical insurance company wants to predict their customers who are going to do renewal in next year, and want you to develop a model with proper EDA to get some recommendation. Below are the steps (question) which you have to follow to create a predictive model.**
2. **You have to write SAS code to get the answer of the below question.**

Question 1:

1.	Import dataset in the SAS environment and check top 5 record of import dataset (2 Mark)

Question 2:

2.	Check data type and dimension of the import dataset (2 Mark)

Question 3:

3.	Check for outlier, if yes then do treatment? (4 Mark)

Question 4:

4.	Checks if any variables have missing values, if yes then do treatment? (4 Mark)

Question 5:

5.	Check percentile distribution summary of all numerical variables? (6 Marks)

Question 6:

6.	Calculate percentage contribution of each class in respective class variables? (6 Marks)

Question 7:

7.	For Xsell purpose sales team wants you to build a macro where they will put the policy number and they will get mobile number and existing premium amount (6 Marks)

Question 8:

8.	Check correlation of all numerical variables to avoid correlated variables in model? (6 Marks)

Question 9:

9.	Create train and test (70:30) dataset from the existing data set. Put seed 1234? (4 Marks)

Question 11:

11.	Create basic logistic model only on numeric variable (no tuning required) for target variables? (3.5 Marks)

Question 12:

10.	Predict test dataset using created model? (3.5 Marks)

Question 13:

10.	Select 0.20 probability as a cutoff and create binary predicted variable. Create confusion matrix on the calculated variable and target variable? (3 Marks)