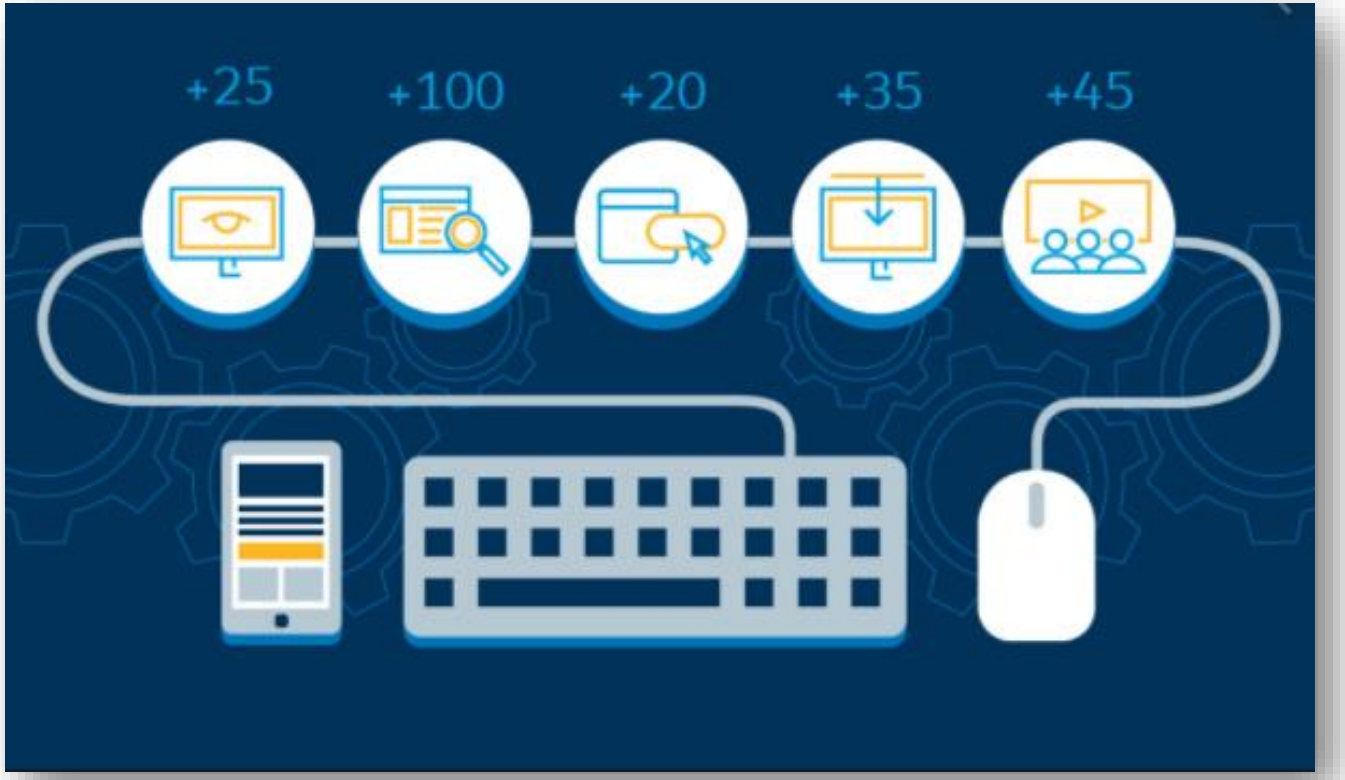
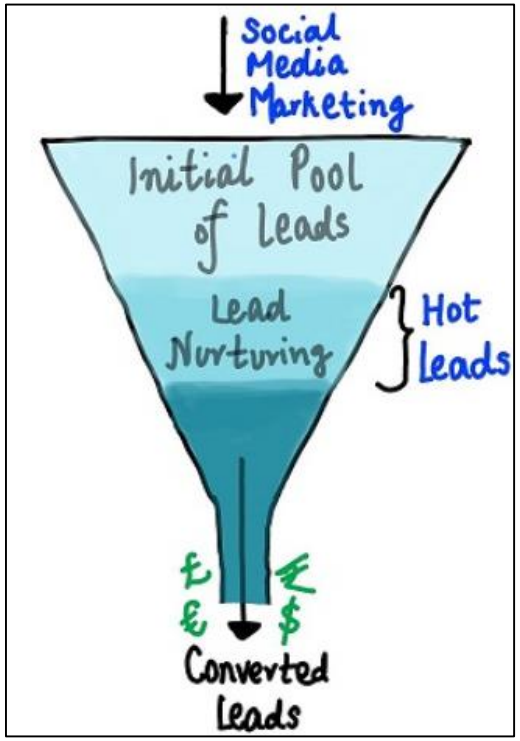
Lead Scoring Presentation



- Sagnik Ghosh

Problem Statement

- An education company named X Education sells online courses to industry professionals. The company markets its courses on several websites and search engines like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Now, although X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted.



Objectives

- X Education has appointed you to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers.

- The company requires you to build a model wherein you need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.

Analysis Approach

- **Cleaning data:** Start off with the necessary data inspection and data cleaning. We need to check shape, info, description, dtypes of the feature columns. Drop columns that are having high percentage of missing values (cut off taken as 45%), having skewed data. Perform imputation for mission values.

- **Data Imbalance Check:** We must perform performed a data imbalance check on Target column.

- **Exploratory Data Analysis (EDA):** We need to perform univariate and bivariate analysis on both numerical and categorical columns with respect to target column and check the dependencies. Need to analyze outliers among the numerical features.

- **Data Preparation:** Treat the outliers, perform capping to reduce the impact of outlier on model. Create dummies for all categorical features. Perform scaling on train dataset for better interpretation of model regression coefficients.

- **Perform Modelling:** We need to perform feature selection using automated technique like Recursive Feature Elimination (RFE) and manual selection approach using p-value and VIF. Need to build a model with at most 15 features which will give a ballpark of the target lead conversion rate to be around 80%.

- Finally we need to generate a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads.

Data Cleaning

- **Handling 'Select' values in some columns:** Below columns are having 'Select' as one of the category. This is most probably because the person has not filled that field. We will replace this field with NaN

  - ✓ Specialization
  - ✓ How did you hear about X Education
  - ✓ Lead Profile
  - ✓ City

- **Duplicate analysis:** There are no duplicate data present with respect to Unique Identifier columns, **Prospect ID** and **Lead Number**. As both the Prospect ID and Lead number are unique columns that are just indicative of the ID number of the Contacted People, we are dropping these two columns.

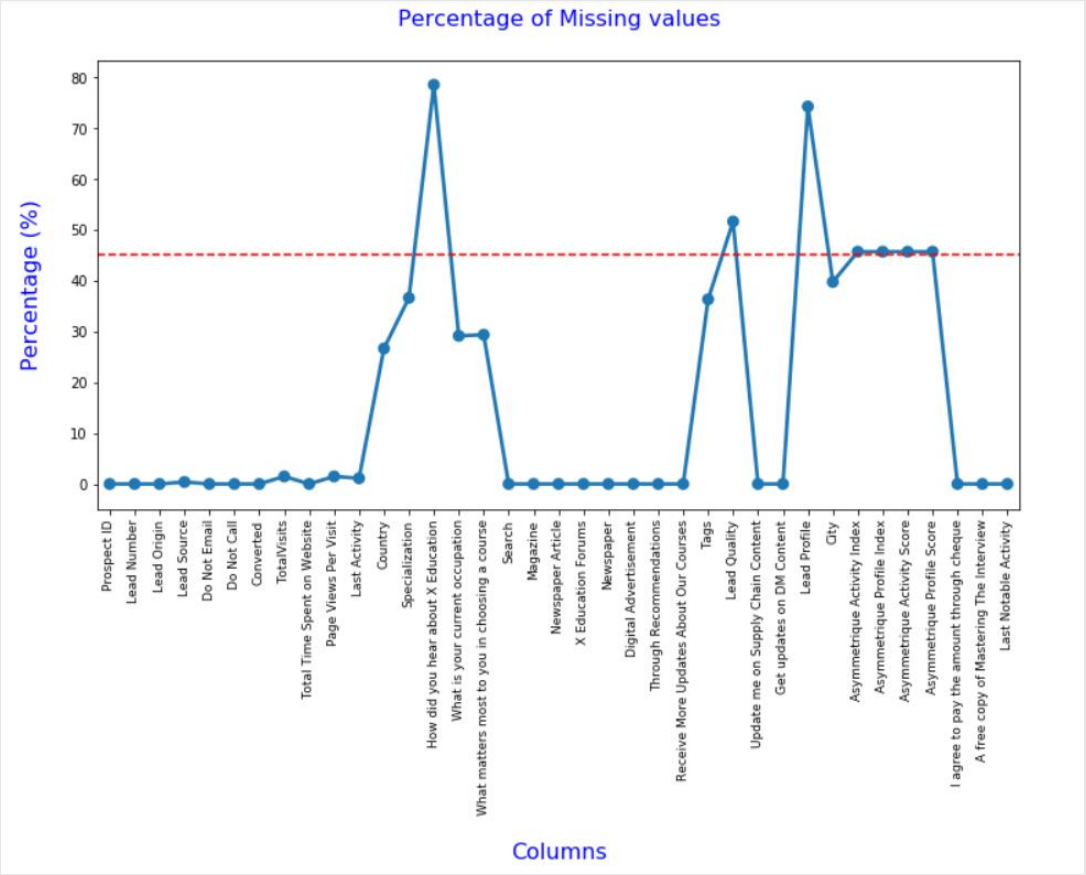
Redundancy Check

- The **Lead Source** column contains one redundant value that is 'Google' and 'google' , we have merged it into one value to 'Google'.

Data Cleaning

Null Value Analysis

- **Column-wise Null value Analysis:** There are **17 columns** with null values. **7 columns** have more than **45% NaN** (null values) which we have dropped as imputing these columns will introduce bias.



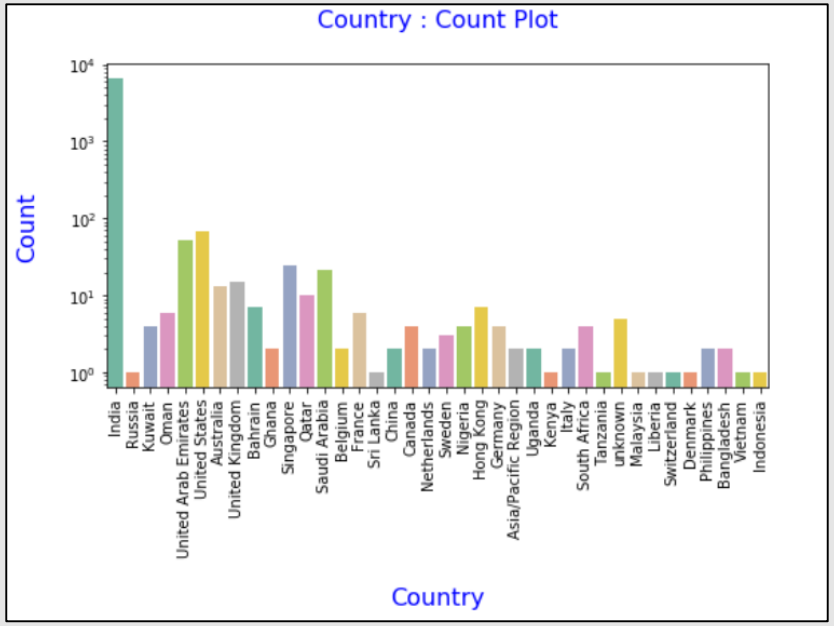
- **Row-wise Null value Analysis:** No rows present which have more than 50% null values.

Data Cleaning

Unnecessary Columns

- **Unnecessary Columns:**
  - ✓ As there is almost 40% unknown values in **City** column, we cannot impute with mode as it is make the whole data skewed. Also, X-Education is online teaching platform. The city information will not be much useful as potential students can available any courses online despite their city.
  - ✓ Most of the information is also erroneous in the city column based on the country.
  - ✓ We have dropped the column from analysis based on above information.

Country	City	City
Australia	Mumbai	6
	Other Cities	2
	Thane & Outskirts	3
Bahrain	Mumbai	1
	Other Cities	2
	Other Cities of Maharashtra	1
	Thane & Outskirts	2
	Tier II Cities	1
Bangladesh	Other Cities	2
Belgium	Mumbai	1
	Thane & Outskirts	1
Canada	Mumbai	3
China	Mumbai	1
Denmark	Other Cities	1
France	Other Cities	2



- ✓ **Country** data is heavily skewed as 95% of the data is mapped as India. Similar to City, Country data is not required for Model building as X-Education is online platform. We have dropped the country columns too.

Data Cleaning

Unnecessary Columns

- There are some variables created by the sales team once they contact the potential lead. These variables will not be available for the model building as these features would not be available before the lead is being contacted. Hence we have dropped these columns.

  - ✓ Lead Profile
  - ✓ Lead Quality
  - ✓ Asymmetrique Profile Score
  - ✓ Asymmetrique Activity Score
  - ✓ Asymmetrique Activity Index
  - ✓ Asymmetrique Profile Index
  - ✓ Tags
  - ✓ Last Notable Activity
  - ✓ Last Activity

Data Cleaning

Skewed Columns

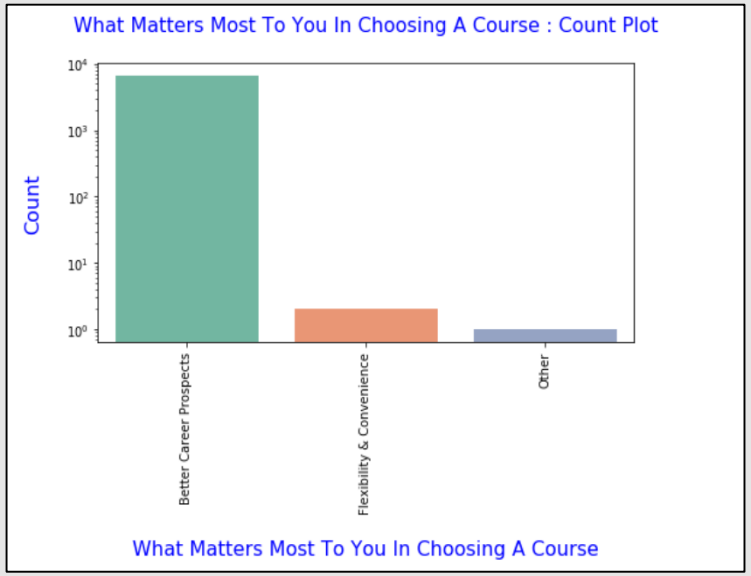
- Some of the columns have only 1 category. These columns will not add any value to the model as the data is skewed. Hence, we have deleted below features.

Column Name	No. of Unique values	No. of Null values	Null Percentage
City	6	3669	39.71
Specialization	18	3380	36.58
What matters most to you in choosing a course	3	2709	29.32
What is your current occupation	6	2690	29.11
Country	38	2461	26.63
Lead Source	20	36	0.39
Lead Origin	5	0	0.00
Digital Advertisement	2	0	0.00
I agree to pay the amount through cheque	1	0	0.00
Get updates on DM Content	1	0	0.00
Update me on Supply Chain Content	1	0	0.00
Receive More Updates About Our Courses	1	0	0.00
Through Recommendations	2	0	0.00
Newspaper Article	2	0	0.00
Newspaper	2	0	0.00
X Education Forums	2	0	0.00
Magazine	1	0	0.00
Search	2	0	0.00
Do Not Call	2	0	0.00

- ✓ I agree to pay the amount through cheque
- ✓ Get updates on DM Content
- ✓ Update me on Supply Chain Content
- ✓ Receive More Updates About Our Courses
- ✓ Magazine

- Below column is having 99% data for one category and very low percentage (0.03 %) of data in other categories. Hence, the data is skewed for below column we have deleted the above column.

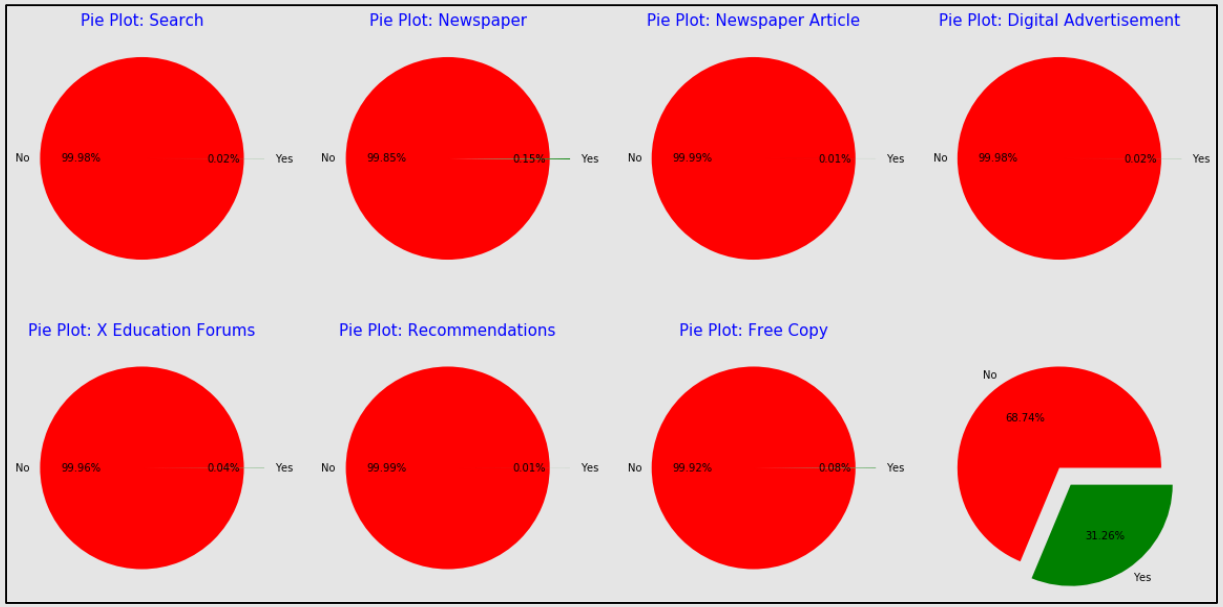
Better Career Prospects	99.954065
Flexibility & Convenience	0.030623
Other	0.015312



Data Cleaning

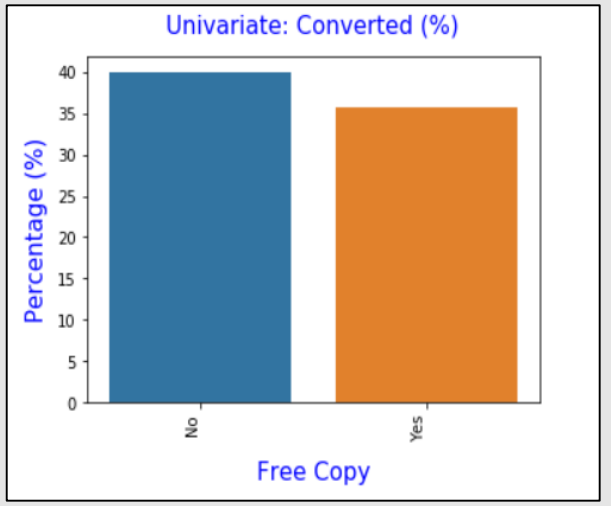
Skewed Columns

- **Search**, **Newspaper**, **Newspaper Article**, **Digital Advertisement, X Education Forums, Recommendation** data are very skewed and can be deleted as they will not add any value to the model.



- Distributing **Free-Copy** of Mastering Interview doesn't seem to add much value as the conversion rate is almost same.

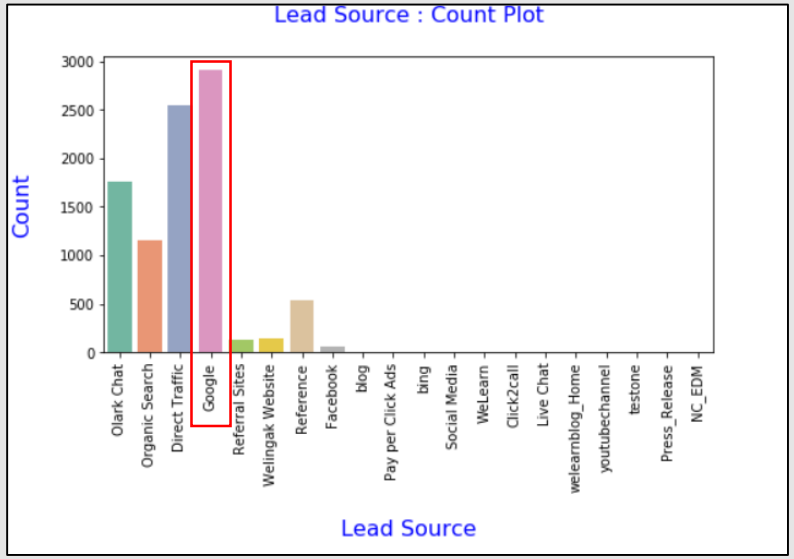
- We have dropped all above columns based on the mentioned facts.



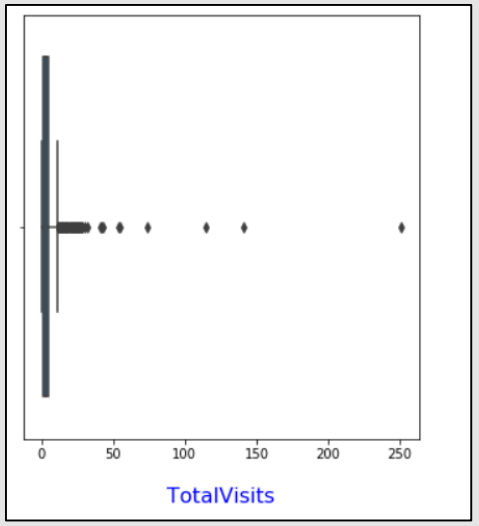
Data Cleaning

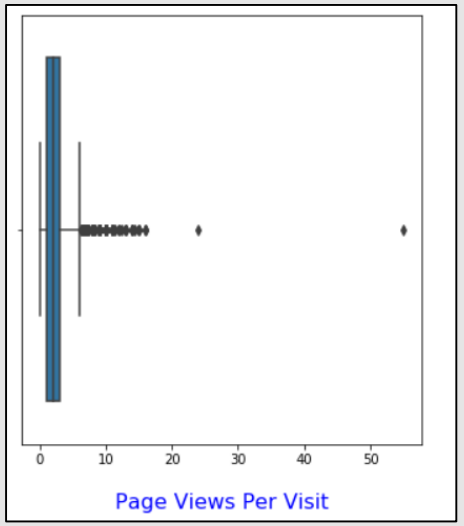
Null value Treatment

Imputing NULL values with Mode : The columns ***Lead Source*** is a categorical variable with some null values. Also majority of the records belong to the Lead Source **Google**. Thus imputed the null values (NaN) for this with mode (most occurring value).



- **Imputing NULL values with Median:** The columns ***TotalVisits*** and ***Page Views Per Visit*** are continuous variables with outliers. Hence the null values for these columns were imputed with the column median values.



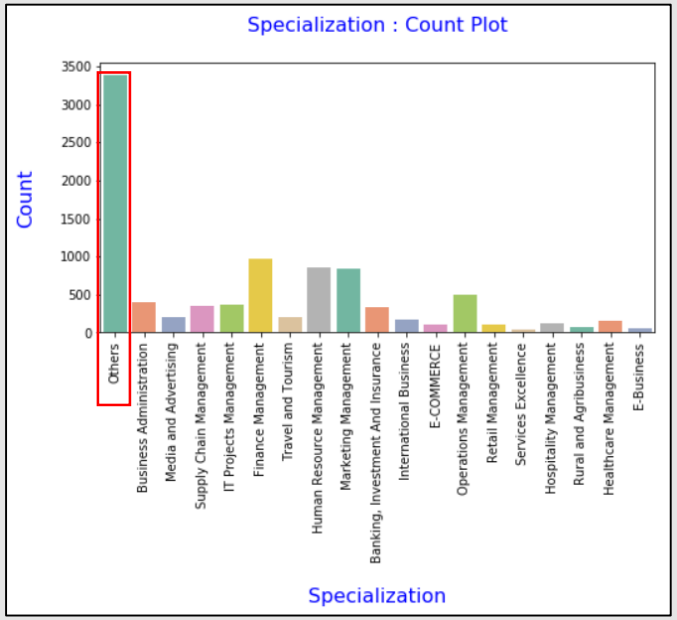


Data Cleaning

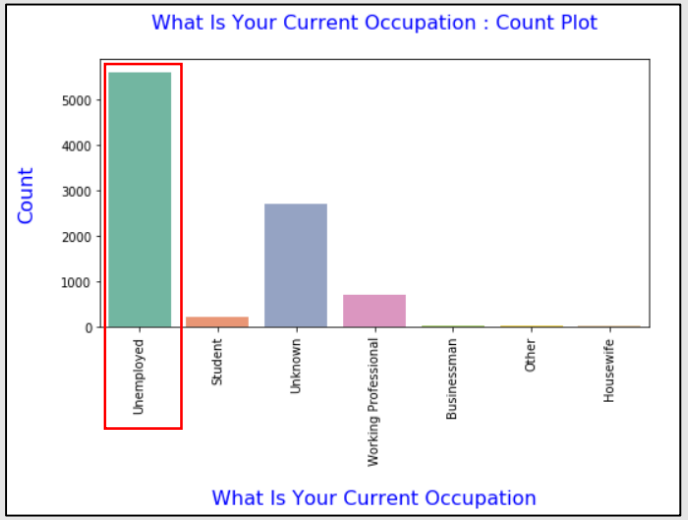
Null value Treatment

- **Imputing NULL values with new category:**

✓ ***Specialization*** column is having 36.58% **NULL** value. It may be possible that the lead has no specialization or may be a student and has no work experience yet , thus he/she has not entered any value. We will create a new category called **Others** to replace the null (NaN) values.



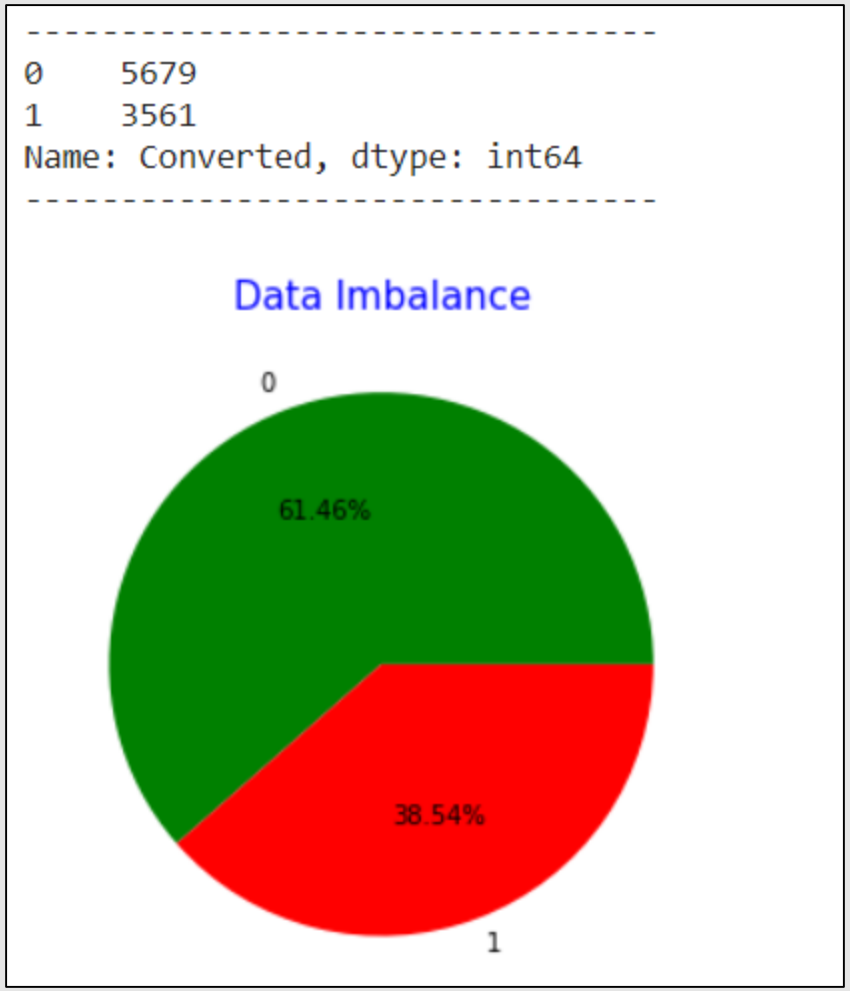
✓ ***What is your current occupation*** column is having 29.11% **NULL** value. Most of the data values are "Unemployed". If we impute the data as "Unemployed" then data will become more skewed. Thus, we will impute the value as **Unknown**.



Data Imbalance

Data distribution in Target Column

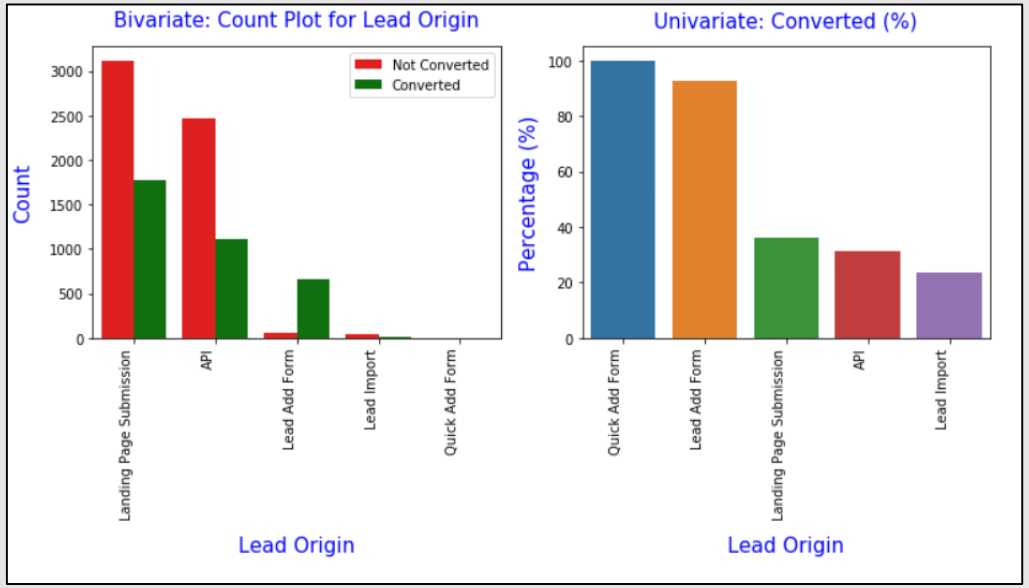
```
---------------------------------
0    5679
1    3561
Name: Converted, dtype: int64
---------------------------------
```



- In the lead conversion ration, 38.54% has converted to leads where as 61.46% did not convert to a lead. So it seems like a almost balanced dataset.

Univariate/Bivariate Analysis

Lead Origin

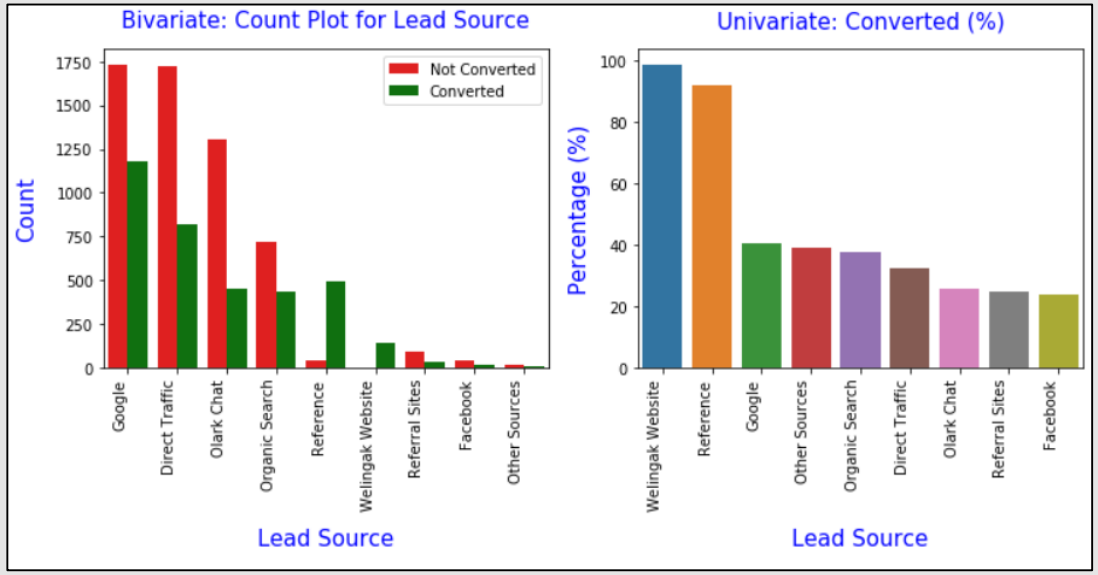


Lead Origin	Converted
Quick Add Form	100.0%
Lead Add Form	92.48%
Landing Page Submission	36.19%
API	31.15%
Lead Import	23.64%

- Most Leads originated from submissions on the landing page and around 36.19% of those are converted followed by API, where around 31.15% are converted.
- Even though Lead Origins from Quick Add Form are 100% Converted, there was just 1 lead from that category.
- Leads from the Lead Add Form are the next highest conversions in this category at around 92.48%.
- Lead Import are very less in count and conversion rate is also the lowest
- To improve overall lead conversion rate, we need to focus more on improving lead conversion of API and Landing Page Submission origin and generate more leads from Lead Add Form.

Univariate/Bivariate Analysis

Lead Source

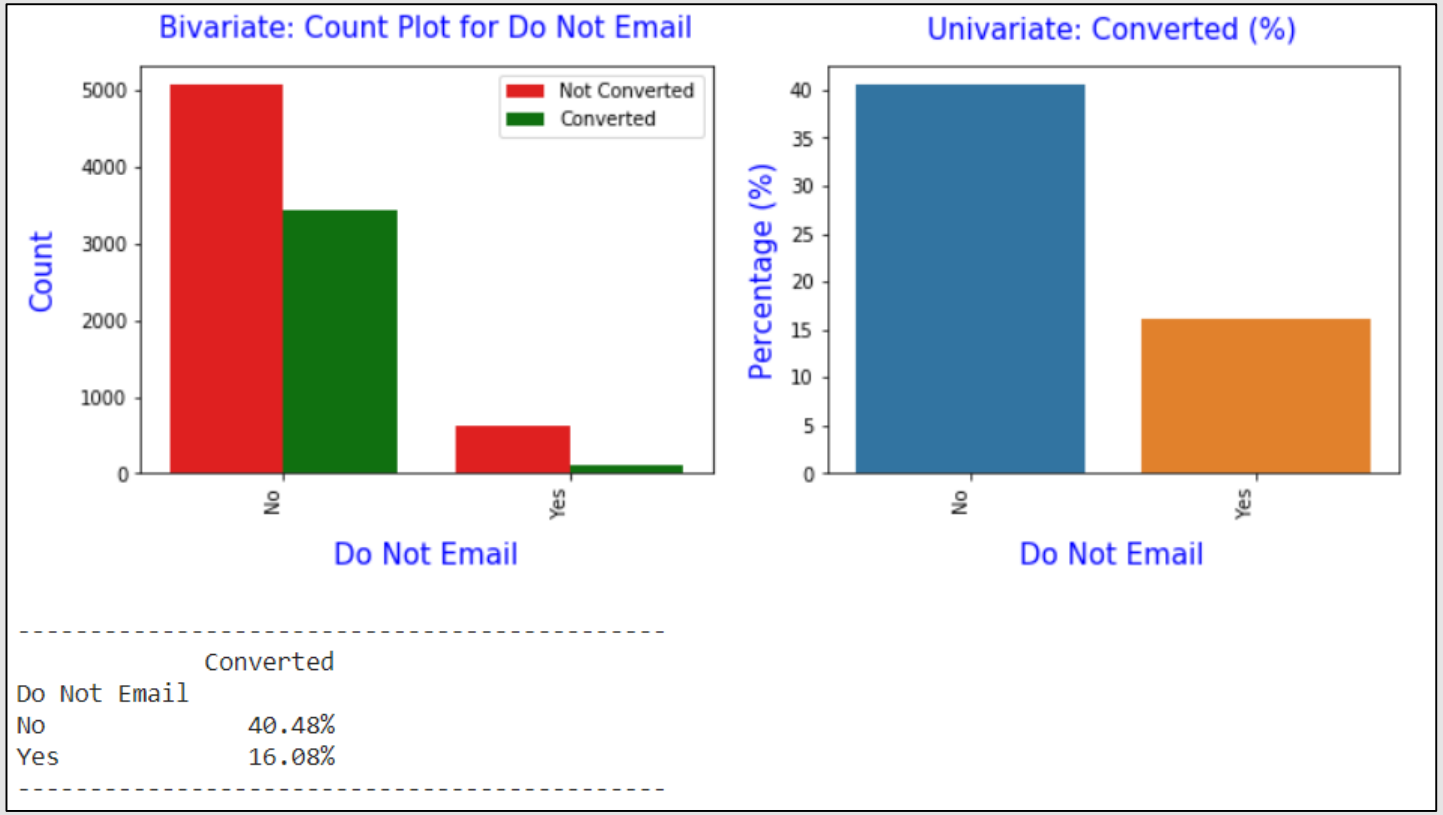


Lead Source	Converted
Welingak Website	98.59%
Reference	91.76%
Google	40.43%
Other Sources	39.13%
Organic Search	37.78%
Direct Traffic	32.17%
Olark Chat	25.53%
Referral Sites	24.8%
Facebook	23.64%

- We have combined smaller lead sources like **Click2call**, **Live Chat**, **NC_EDM** etc. as **Other Sources**.
- The source of the most leads was Google, and 40.43% of the leads converted, followed by Direct Traffic, Organic search and Olark chat where around 32.17%, 37.78% and 25.53% converted respectively.
- A lead that came from a reference has over 91.76% conversion.
- Welingak Website has almost 98.59% lead conversion rate. This option should be explored more to increase lead conversion
- To increase lead count, **initiatives should be taken so already existing members increase their referrals.**

Univariate/Bivariate Analysis

Do Not Email

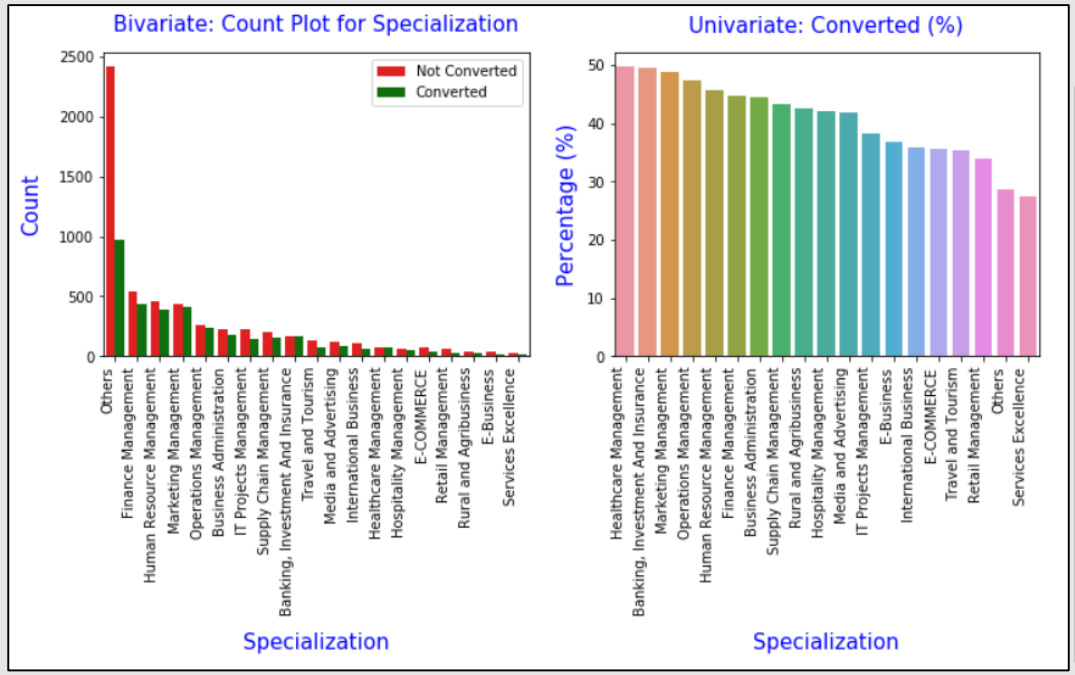


Do Not Email	Converted
No	40.48%
Yes	16.08%

- Majority of the people are ok with receiving email (~92%)
- People who are ok with email has conversion rate of 40.48%
- People who have opted out of receive email has lower rate of conversion (only 16.08%)

Univariate/Bivariate Analysis

Specialization

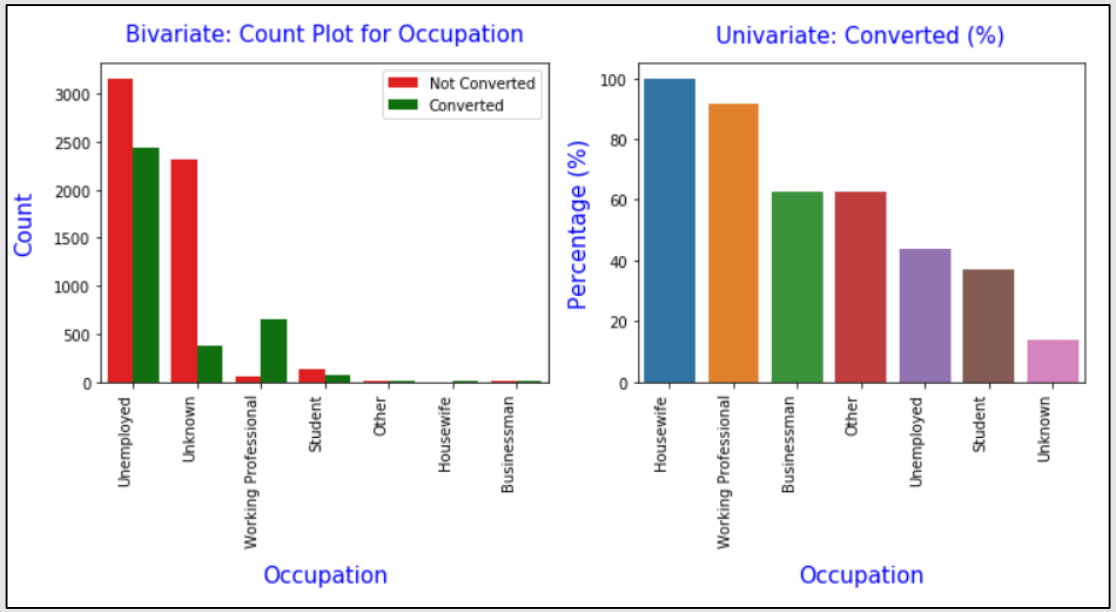


Specialization	Converted
Healthcare Management	49.69%
Banking, Investment And Insurance	49.41%
Marketing Management	48.69%
Operations Management	47.32%
Human Resource Management	45.75%
Finance Management	44.67%
Business Administration	44.42%
Supply Chain Management	43.27%
Rural and Agribusiness	42.47%
Hospitality Management	42.11%
Media and Advertising	41.87%
IT Projects Management	38.25%
E-Business	36.84%
International Business	35.96%
E-COMMERCE	35.71%
Travel and Tourism	35.47%
Retail Management	34.0%
Others	28.67%
Services Excellence	27.5%

- Most of the leads have not mentioned a specialization and around 28.67% of those converted
- Leads with Healthcare Management, Operations Management, Human Resource Management, Finance management and Marketing Management - Over 45% Converted

Univariate/Bivariate Analysis

Occupation

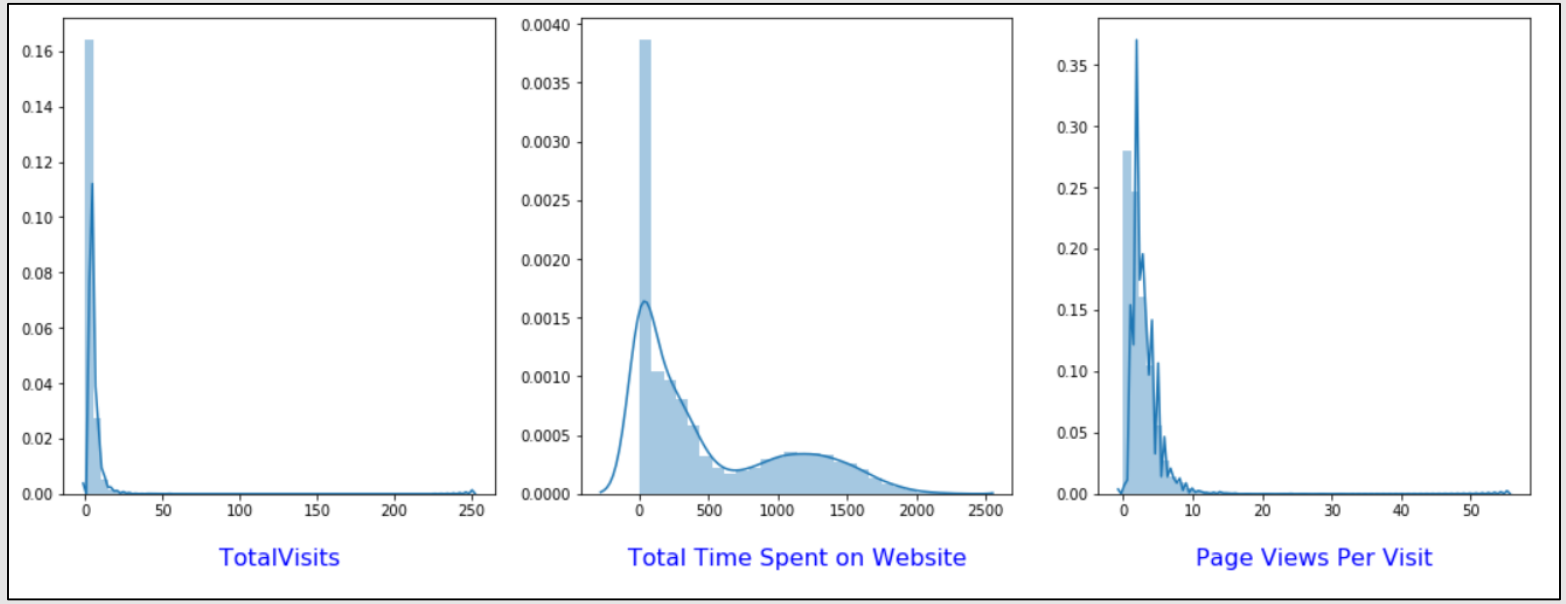


Occupation	Converted
Housewife	100.0%
Working Professional	91.64%
Businessman	62.5%
Other	62.5%
Unemployed	43.59%
Student	37.14%
Unknown	13.75%

- Though Housewives are less in numbers, they have 100% conversion rate
- Working professionals, Businessmen and Other category have high conversion rate
- Though Unemployed people have been contacted in the highest number, the conversion rate is low (43.59%) We cannot combine lower value categories like Unknown, Other as their conversion rate is very different. Combing them may provide wrong predictions.

Univariate Analysis

Numerical Features



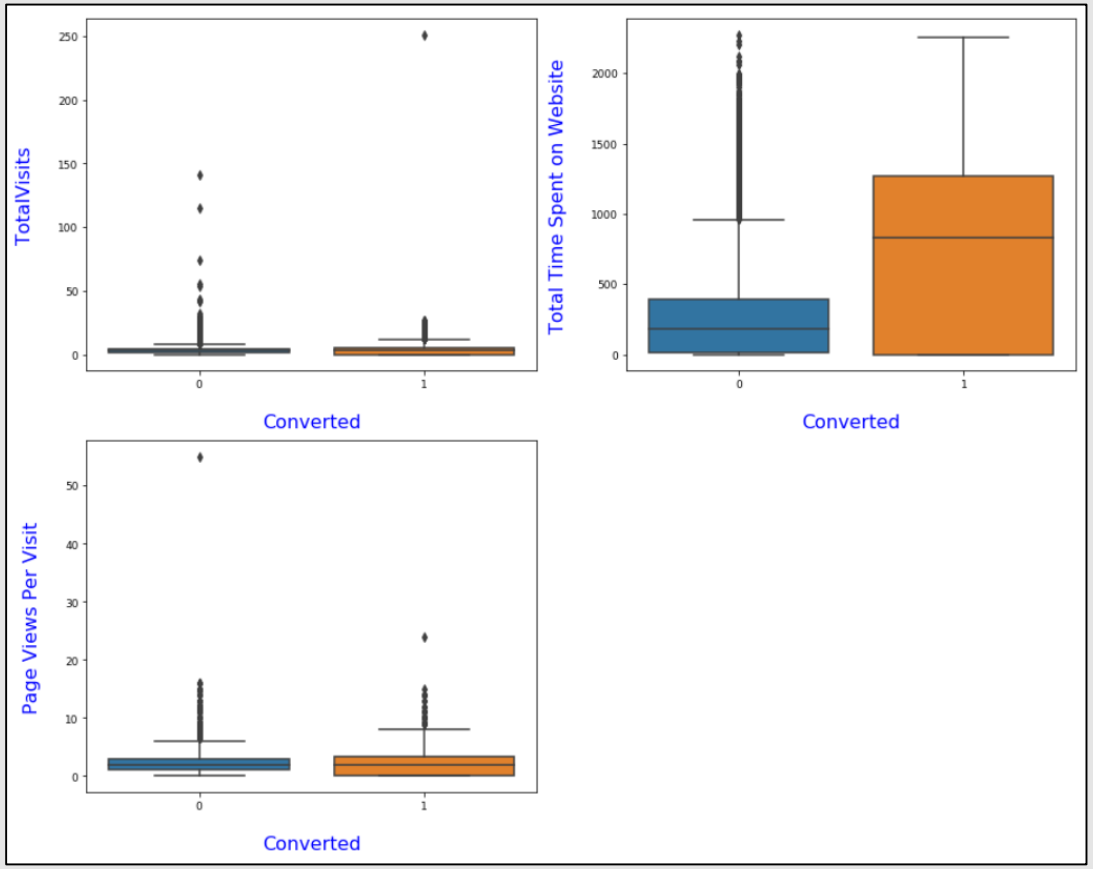
- Data on Total Visits , Page Views per Visit and Total Time Spent on Website columns are not normally distributed and seems to be skewed.

Bivariate Analysis

Converted vs Total Visits

Converted vs Total Time Spent on Website

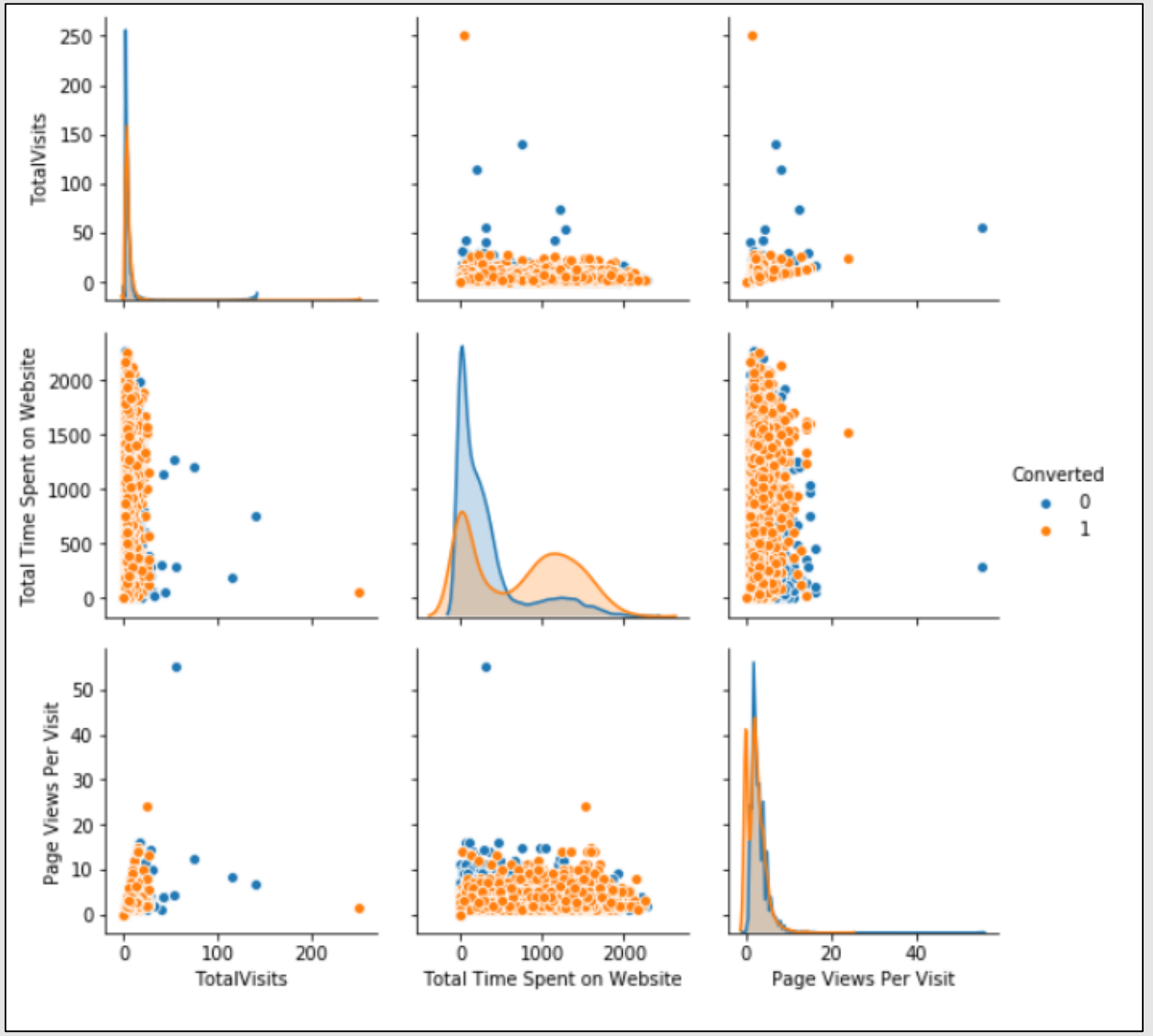
Converted vs Page Views Per Visit



- Total Visits ,Page Views per Visit and Total Time Spent on Website has some outliers which needs to be treated.
- Total Time Spent on Website column has highest conversion rate , followed by Page Views per Visit.
- Total Visits has lowest conversion rate.

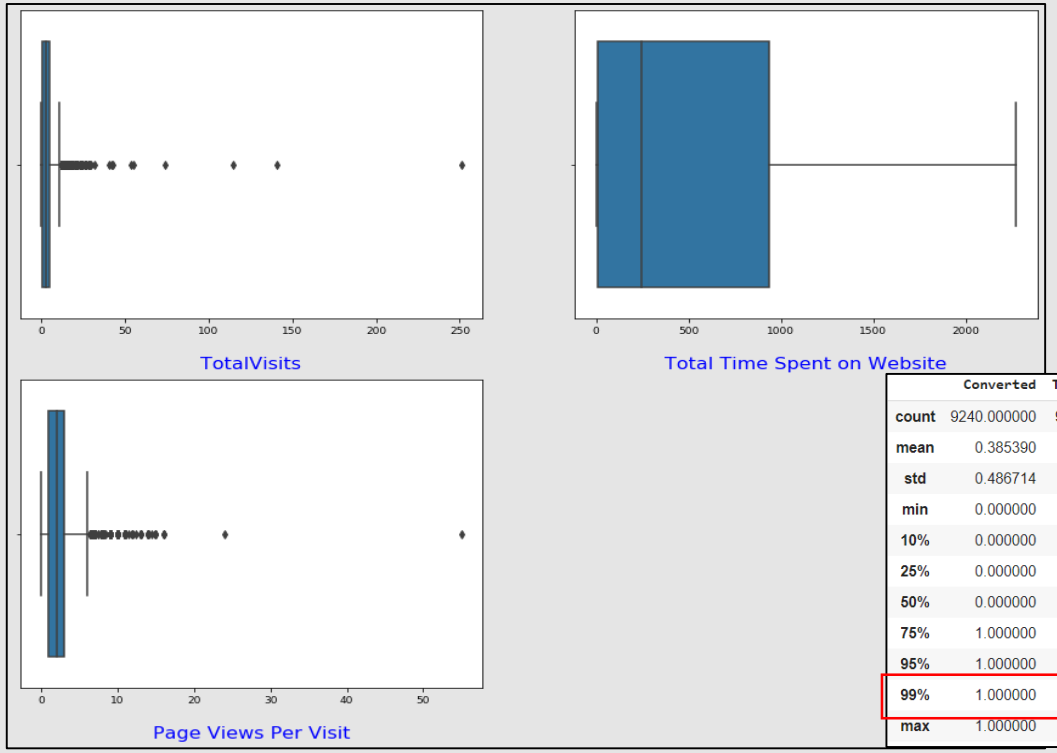
Bivariate Analysis

Numerical vs Numerical



Data Preparation

Outlier Treatment



	Converted	TotalVisits	Total Time Spent on Website	Page Views Per Visit
count	9240.000000	9240.000000	9240.000000	9240.000000
mean	0.385390	3.438636	487.698268	2.357440
std	0.486714	4.819024	548.021466	2.145781
min	0.000000	0.000000	0.000000	0.000000
10%	0.000000	0.000000	0.000000	0.000000
25%	0.000000	1.000000	12.000000	1.000000
50%	0.000000	3.000000	248.000000	2.000000
75%	1.000000	5.000000	936.000000	3.000000
95%	1.000000	10.000000	1562.000000	6.000000
99%	1.000000	17.000000	1840.610000	9.000000
max	1.000000	251.000000	2272.000000	55.000000

Though outliers in Total Visits and Page Views Per Visit shows valid values, this will misclassify the outcomes and consequently create problems when making inferences with the wrong model. Logistic Regression is influenced by outliers. So lets cap the Total Visits and Page Views Per Visit to their 95th percentile due to following reasons:

- Data set is fairly high number
- 95th percentile and 99th percentile of these columns are very close and hence impact of the capping to 95th or 99th percentile will be the same.

Creating Dummies

Binary Encoding & Categorical to Dummy Variable conversion

Do Not Email

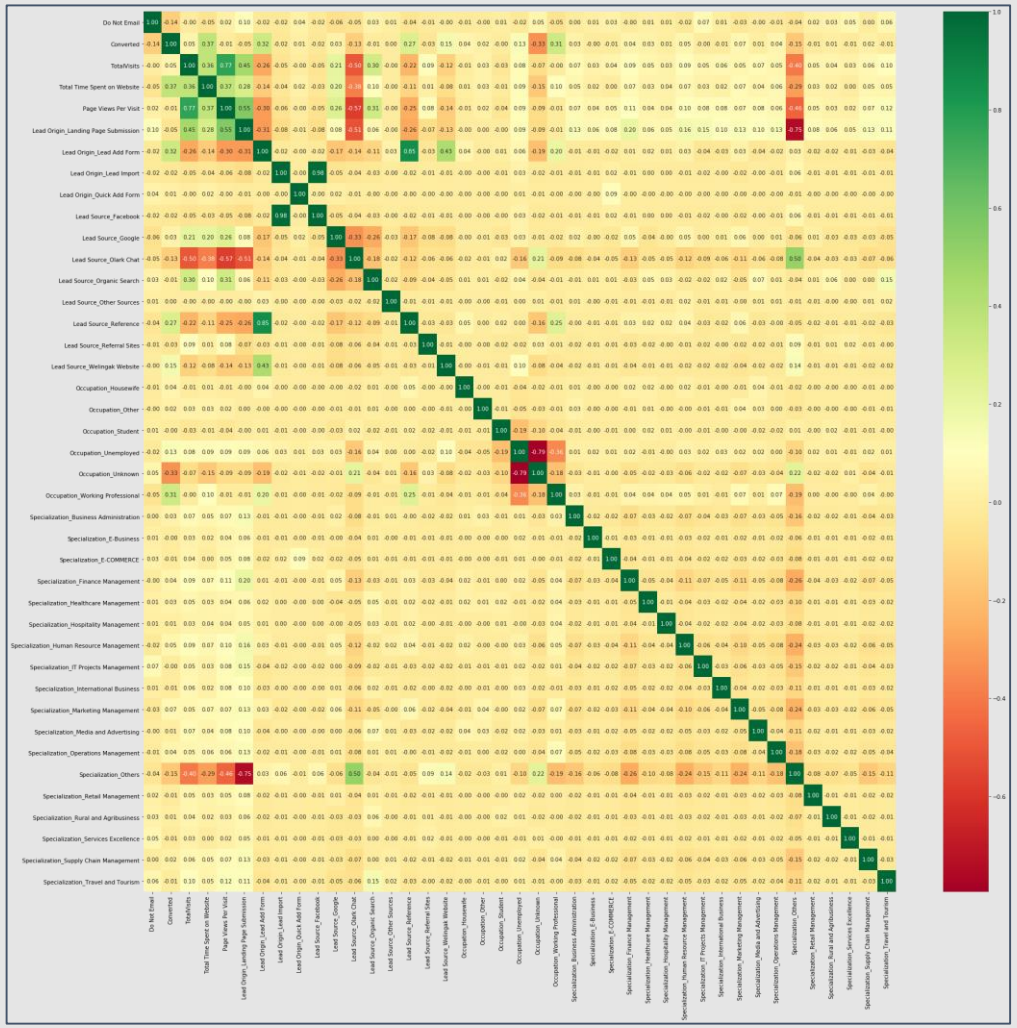
- Converting the binary variables (Yes/No) to 0/1.

Lead Origin, Lead Source, Occupation, Specialization

- For the categorical variables with multiple levels, dummy features (one-hot encoded) were created.

Correlation matrix

Checking the correlations between columns



- We have dropped columns Occupation_Unknown, Lead Source_Reference and Lead Source_Facebook as these columns has high collinearity with other columns (> 0.75) and will effect the VIF (variance inflation factor) due to multicollinearity in the features. Hence we have dropped it:

Feature engineering

Train Test split & Scaling

Train Test Split

- The original data frame was split into train and test dataset.
- The train dataset was used to train the model and test dataset was used to evaluate the model.

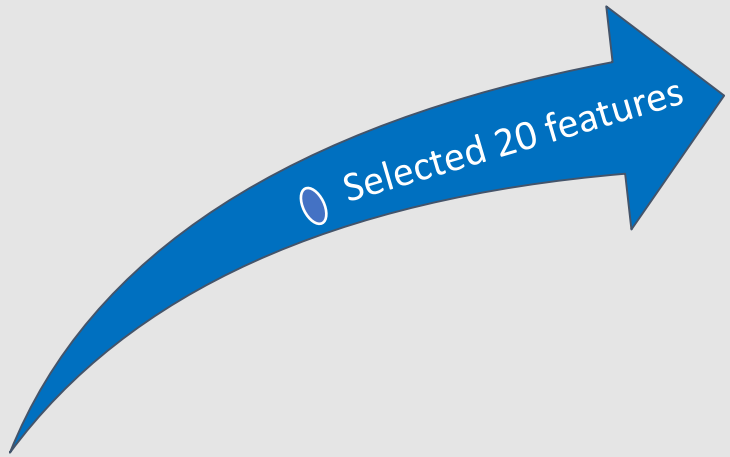
Scaling

- Scaling helps in interpretation. It is important to have all variables scale free
- 'Standardisation' was used to scale the data for modelling. It basically brings all of the data into a standard normal distribution with mean at zero and standard deviation one.

Feature Selection

RFE

- **Recursive feature elimination :** It is an optimization technique for finding the best performing subset of features. It is based on the idea of repeatedly constructing a model and choosing either the best (based on coefficients), setting the feature aside and then repeating the process with the rest of the features. This process is applied until all the features in the dataset are exhausted. Features are then ranked according to when they were eliminated.




```
['Do Not Email',
 'Total Time Spent on Website',
 'Lead Origin_Landing Page Submission',
 'Lead Origin_Lead Add Form',
 'Lead Origin_Lead Import',
 'Lead Source_Google',
 'Lead Source_Olark Chat',
 'Lead Source_Other Sources',
 'Lead Source_Referral Sites',
 'Lead Source_Welingak Website',
 'Occupation_Housewife',
 'Occupation_Student',
 'Occupation_Unemployed',
 'Occupation_Working Professional',
 'Specialization_E-COMMERCE',
 'Specialization_Hospitality Management',
 'Specialization_International Business',
 'Specialization_Others',
 'Specialization_Retail Management',
 'Specialization_Rural and Agribusiness']
```

Model Building

p-value & VIF (Model Summary)

- Generalized Linear Models from Stats Models is used to build the Logistic Regression model.
- The model is built initially with the **20 variables selected** by RFE.
- Unwanted features are dropped serially after checking **p values (<0.05) and VIF (< 5)** and model is built multiple times.
- The final model with **12 features**, passes both the significance test and the multi-collinearity test.

Generalized Linear Model Regression Results

Dep. Variable:	Converted	**No. Observations:**	6468
Model:	GLM	**Df Residuals:**	6455
Model Family:	Binomial	**Df Model:**	12
Link Function:	logit	**Scale:**	1.0000
Method:	IRLS	**Log-Likelihood:**	-2850.2
Date:	Mon, 07 Dec 2020	**Deviance:**	5700.4
Time:	11:31:12	**Pearson chi2:**	7.91e+03
No. Iterations:	7		
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
const	-1.2912	0.147	-8.778	0.000	-1.579	-1.003
Do Not Email	-1.2755	0.160	-7.948	0.000	-1.590	-0.961
Total Time Spent on Website	1.0901	0.038	28.609	0.000	1.015	1.165
Lead Origin_Landing Page Submission	-0.7466	0.124	-6.015	0.000	-0.990	-0.503
Lead Origin_Lead Add Form	3.3560	0.203	16.496	0.000	2.957	3.755
Lead Source_Google	0.2569	0.077	3.327	0.001	0.106	0.408
Lead Source_Olark Chat	1.0966	0.125	8.794	0.000	0.852	1.341
Lead Source_Welingak Website	2.5025	0.743	3.368	0.001	1.046	3.959
Occupation_Student	1.0501	0.229	4.588	0.000	0.601	1.499
Occupation_Unemployed	1.1512	0.082	13.957	0.000	0.990	1.313
Occupation_Working Professional	3.5509	0.196	18.140	0.000	3.167	3.935
Specialization_Hospitality Management	-0.9297	0.315	-2.951	0.003	-1.547	-0.312
Specialization_Others	-0.9312	0.117	-7.928	0.000	-1.161	-0.701

	Features	VIF
0	Occupation_Unemployed	2.71
1	Specialization_Others	2.42
2	Lead Origin_Landing Page Submission	2.39
3	Lead Source_Olark Chat	2.04
4	Lead Origin_Lead Add Form	1.64
5	Lead Source_Google	1.64
6	Occupation_Working Professional	1.32
7	Lead Source_Welingak Website	1.27
8	Total Time Spent on Website	1.26
9	Do Not Email	1.11
10	Occupation_Student	1.06
11	Specialization_Hospitality Management	1.02

Predicting Model Probability

On Train Dataset

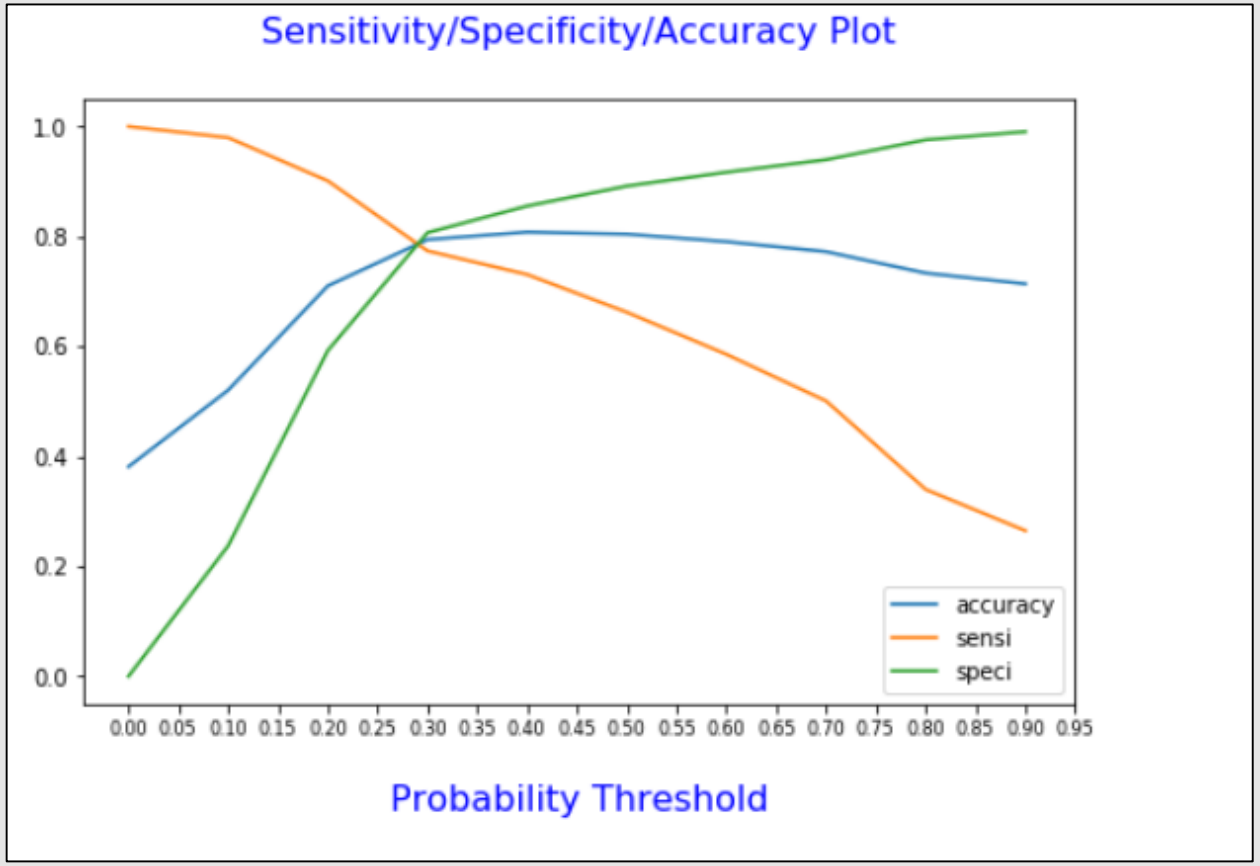
	Converted	Converted_Prob	Prospect_Id
0	0	0.493596	1871
1	0	0.142528	6795
2	0	0.323444	3516
3	0	0.716628	8105
4	0	0.277508	3934

- Creating a data frame with the actual Converted flag and the predicted probabilities. Showing top 5 records of the data frame.

Finding Optimal Probability Threshold

Optimal Cutoff is 0.288

- The **accuracy sensitivity and specificity** was calculated for various values of probability threshold and plotted in the graph to the right.
- From the curve above, **0.288 is found to be the optimum point** for cutoff probability.
- At this threshold value, all the 3 metrics -accuracy sensitivity and specificity was found to be **around 78% which is a well acceptable value**.

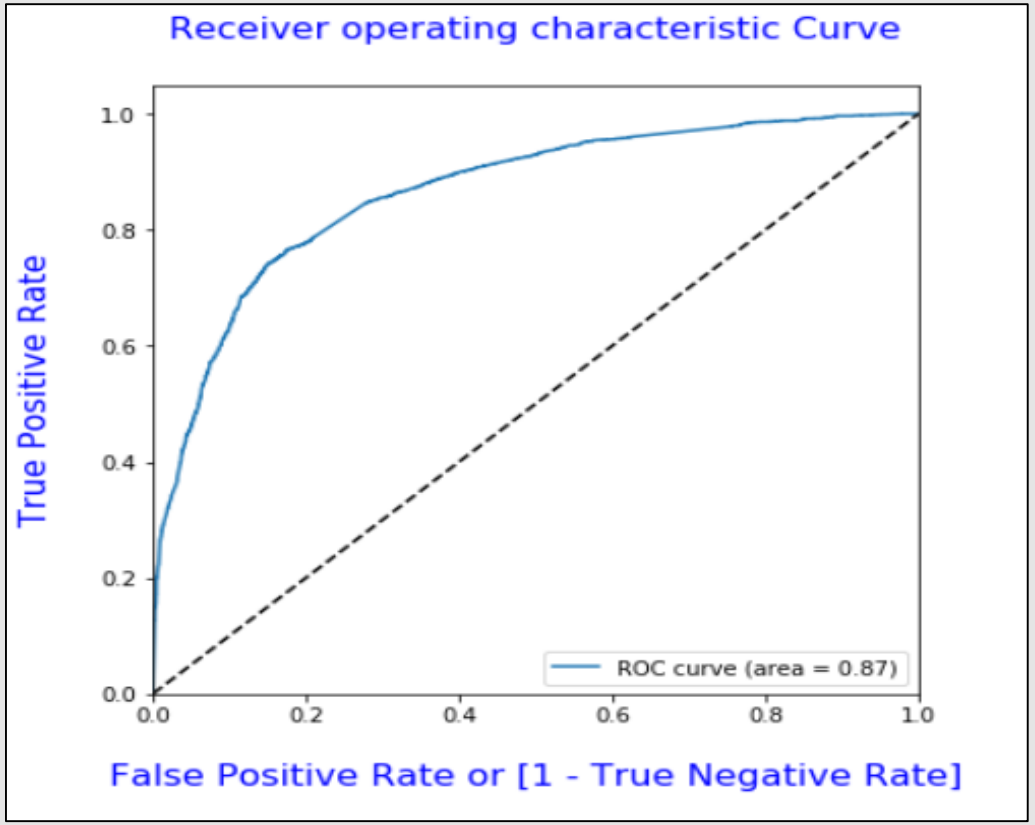


Receiver Operating Characteristics (ROC) Curve

Area under the Curve (GINI) is 0.87

An ROC curve demonstrates below:

- It shows the **tradeoff between sensitivity and specificity** (any increase in sensitivity will be accompanied by a decrease in specificity)
- The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test.
- The closer the curve comes to the **45-degree diagonal of the ROC space**, the less accurate the test.
- The area under the **Curve or Gini is 0.87 which is represent a good model**.



Confusion matrix & Other metrics

Accuracy, Recall, Precision, Specificity and F1 Score on train Dataset

```
array([[3187,  815],
       [ 541, 1925]], dtype=int64)
```

```
---------------------------------------------------------------
The Recall score of the model is 0.7806163828061639
The Specificity score of the model is  0.796351824087956
The Precision score of the model is  0.7025547445255474
The False positive rate of the model is  0.20364817591204398
The Negative predicted value of the model is  0.8548819742489271
---------------------------------------------------------------
```

	Converted	Converted_Prob	Prospect_Id	final_predicted
0	0	0.493596	1871	1
1	0	0.142528	6795	0
2	0	0.323444	3516	1
3	0	0.716628	8105	1
4	0	0.277508	3934	0

- The **Recall/Sensitivity score** of the model **is 78.1%.**Out of actual Converted leads the model has predicted 78% Converted leads correctly.
- The **Specificity score of the model is 79.6%.**Out of actual non Converted leads the model has predicted 79.6% non Converted leads correctly.
- The **Precision score of the model is 70.3%.**Out of predicted Converted leads the model has predicted 70.3% Converted leads correctly.
- Sensitivity/Recall in this case indicates how many leads the model identify correctly out of all potential leads which are converting. Almost around 80% is what the CEO has requested in this case study.
- The **F1 score for the model for train dataset is 74%.**If we give equal importance to Precision and Recall , then we calculate F1 score and see the model F1 score is 74%.

Predicting Model Probability

On Test Dataset

	Converted	Converted_Prob	Prospect_Id
0	1	0.246544	4269
1	1	0.786261	2376
2	1	0.839730	7766
3	0	0.277508	9199
4	1	0.903241	4359

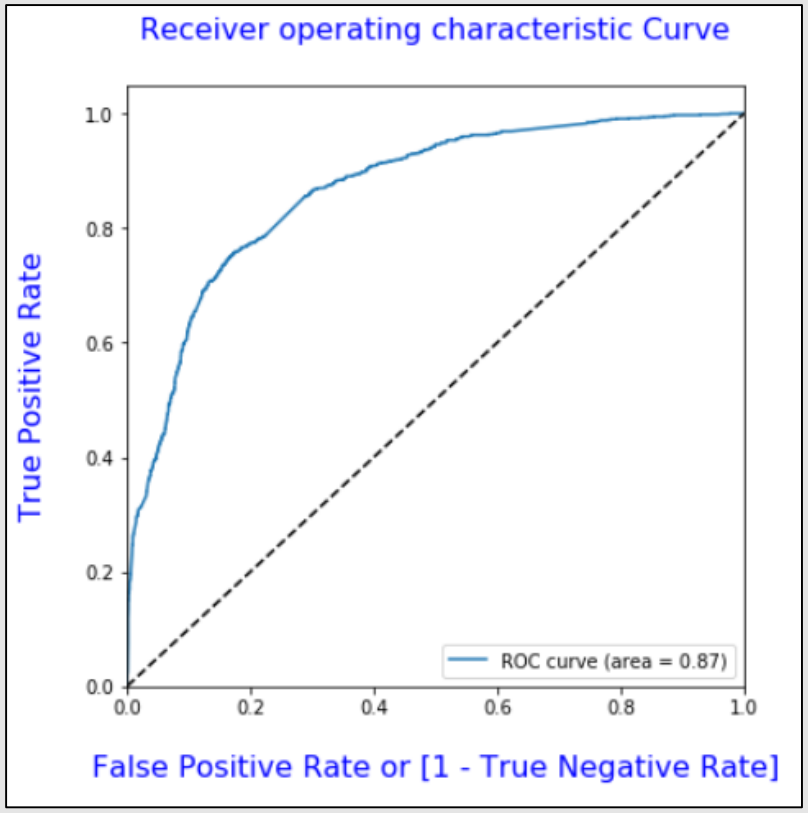
- Creating a data frame with the actual Converted flag and the predicted probabilities. Showing top 5 records of the data frame.
- Predicting the Converted values on Test dataset by taking the threshold of **0.288**

Receiver Operating Characteristics (ROC) Curve

Area under the Curve (GINI) is 0.87 in Test dataset

An ROC curve demonstrates on Test Dataset below:

- It shows the **tradeoff between sensitivity and specificity** (any increase in sensitivity will be accompanied by a decrease in specificity)
- The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test.
- The closer the curve comes to the **45-degree diagonal of the ROC space**, the less accurate the test.
- The area under the **Curve or Gini is 0.87 which is represent a good model**.



Confusion matrix & Other metrics

Accuracy, Recall, Precision, Specificity on test Dataset

```
array([[1317,  360],
       [ 239,  856]], dtype=int64)
```

```
---------------------------------------------------------------
The Recall score of the model is 0.7817351598173516
The Specificity score of the model is  0.7853309481216458
The Precision score of the model is  0.7039473684210527
The False positive rate of the model is  0.2146690518783542
The Negative predicted value of the model is  0.846401028277635
---------------------------------------------------------------
```

	Converted	Converted_Prob	Prospect_Id	final_predicted
0	1	0.246544	4269	0
1	1	0.786261	2376	1
2	1	0.839730	7766	1
3	0	0.277508	9199	0
4	1	0.903241	4359	1

- **The Recall/Sensitivity score of the model is 78.1%.**Out of actual Converted leads the model has predicted 78% Converted leads correctly.
- The **Specificity score of the model is 79%.**Out of actual non Converted leads the model has predicted 79.6% non Converted leads correctly.
- **The Precision score of the model is 70.4%.**Out of predicted Converted leads the model has predicted 70.4% Converted leads correctly.
- Sensitivity/Recall in this case indicates how many leads the model identify correctly out of all potential leads which are converting. Almost around 80% is what the CEO has requested in this case study.

The Sensitivity/Recall value on Test data is 78.2% vs 78.1% in Train data. The accuracy values is 78%. It shows that model is performing well in test data set also and is not over-trained.

Lead Score Generation

On Test dataset

Lead Score is calculated for all the leads in the Test data frame.

Formula for Lead Score calculation is:

Lead Score = 100 * Conversion Probability

	Prospect_Id	Converted	Converted_Prob	final_predicted	Lead_score
546	3478	1	0.999614	1	99
2405	5921	1	0.999443	1	99
224	8120	1	0.999311	1	99
835	4613	1	0.999085	1	99
745	6383	1	0.999020	1	99
1091	818	1	0.999020	1	99
2589	7327	1	0.998810	1	99
2150	133	1	0.998732	1	99
605	7187	1	0.998601	1	99
1242	8107	1	0.998496	1	99

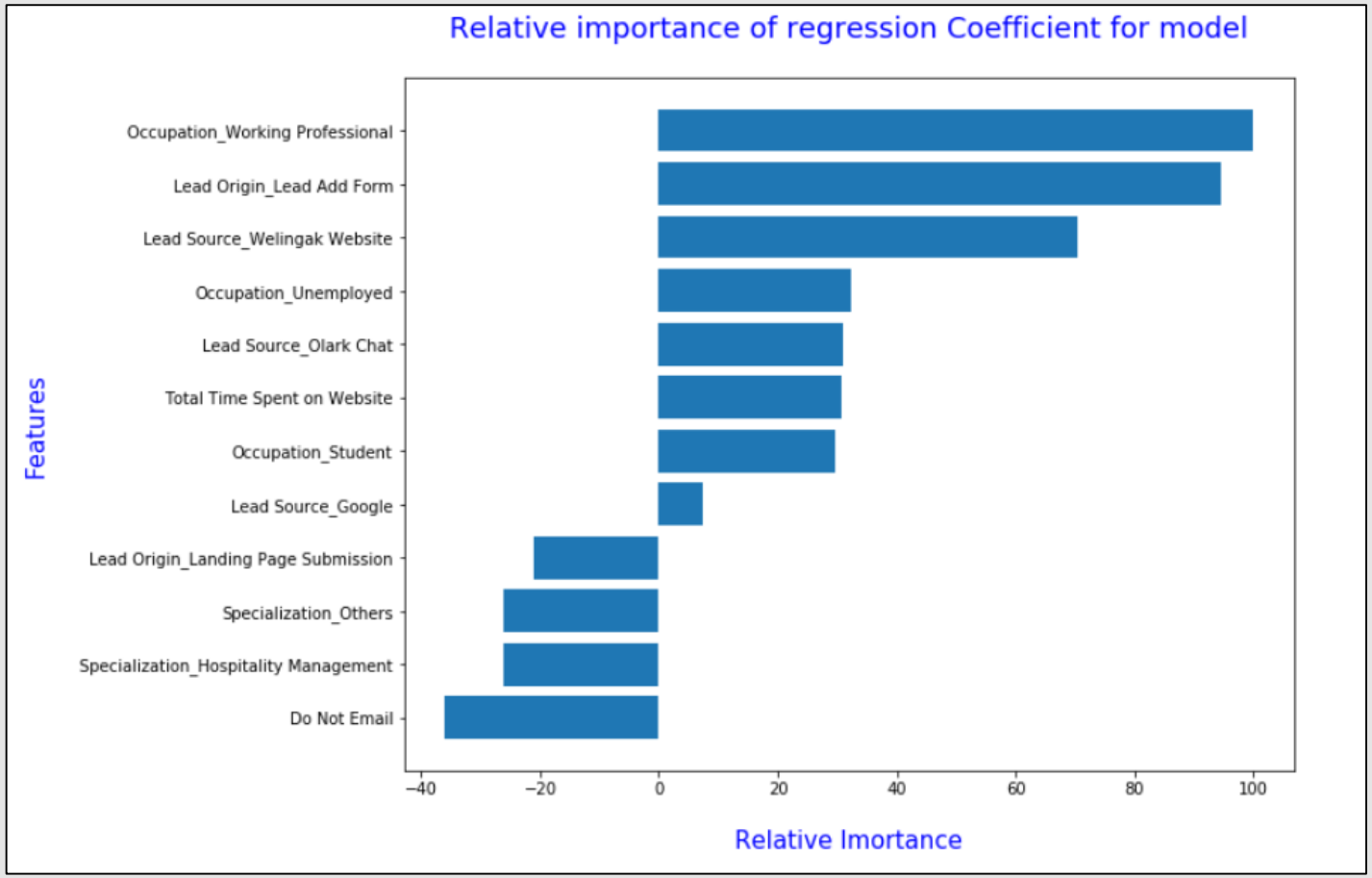
- The Conversion Probability is multiplied by 100 to obtain the Lead Score for each lead.
- Higher the lead score, higher is the probability of a lead getting converted and vice versa,
- Since, we had used 0.288 as our final Probability threshold for deciding if a lead will convert or not, any lead with a lead score of 29 or above will have a value of ‘1’ in the final_predicted column.

Determining Feature Importance

Getting a relative coefficient value for all the features with respect to the feature's with the highest coefficient

Final Features list based on its importance:

- It was found that the below variables/Features mattered the most based on which the leads are most likely to convert into paying customers (In descending order).



Model Summary

The magnitude and sign of the coefficients loaded in the logit function

logit(p) = log(p/(1-p)) = (3.36 * Lead Origin_Lead Add Form) + (3.55 * Occupation_Working Professional) + (2.50 * Lead Source_Welingak Website) + (1.15 * Occupation_Unemployed) + (1.10 * Lead Source_Olark Chat) + (1.09 * Total Time Spent on Website) + (1.05 * Occupation_Student) + (0.26 * Lead Source_Google) - (0.75 * Lead Origin_Landing Page Submission) - (0.93 * Specialization_Hospitality Management) -(0.93 * Specialization_Others) - (1.28 * Do Not Email) - 1.29

In general, we can have multiple predictor variables in a logistic regression model as below:

logit(p) = log(p/(1-p))= β0 + β1* X1 + … + βn * Xn

Applying such a model to our example dataset, each estimated coefficient is the expected change in the log odds of being a potential lead for a unit increase in the corresponding predictor variable holding the other predictor variables constant at a certain value.
Each exponentiated coefficient is the ratio of two odds, or the change in odds in the multiplicative scale for a unit increase in the corresponding predictor variable holding other variables at a certain value.

Point to remember:

Another point to note here is that, depending on the business requirement, we can increase or decrease the probability threshold value with in turn will decrease or increase the Sensitivity and increase or decrease the Specificity of the model.

High Sensitivity will ensure that almost all leads who are likely to Convert are correctly predicted where as high Specificity will ensure that leads that are on the brink of the probability of getting Converted or model predicted as not converted properly when compare to actual not converted leads.

END