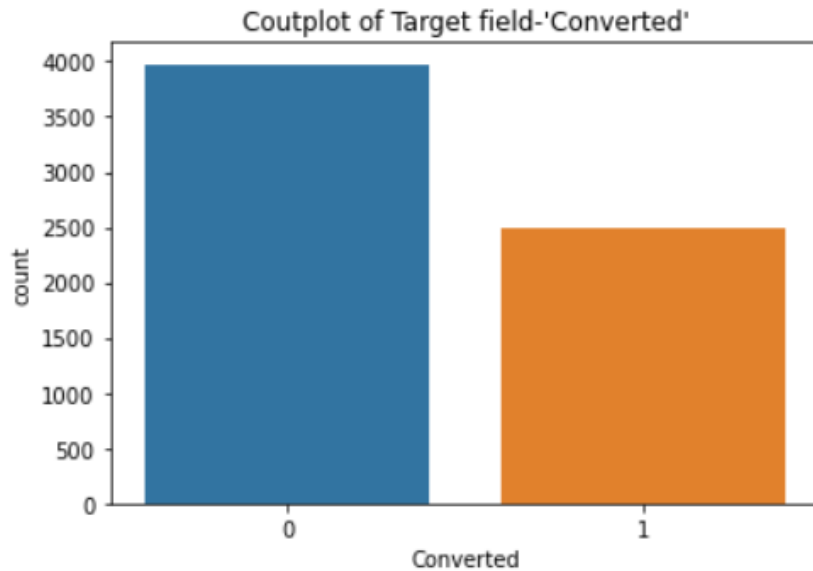


## Presentation

- We have ~40% customers that are likely to be converted.

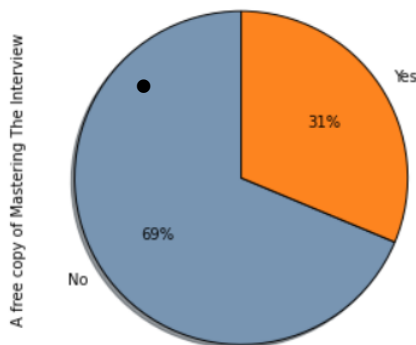


```
] : train_new['Converted'].value_counts(normalize=True)
```

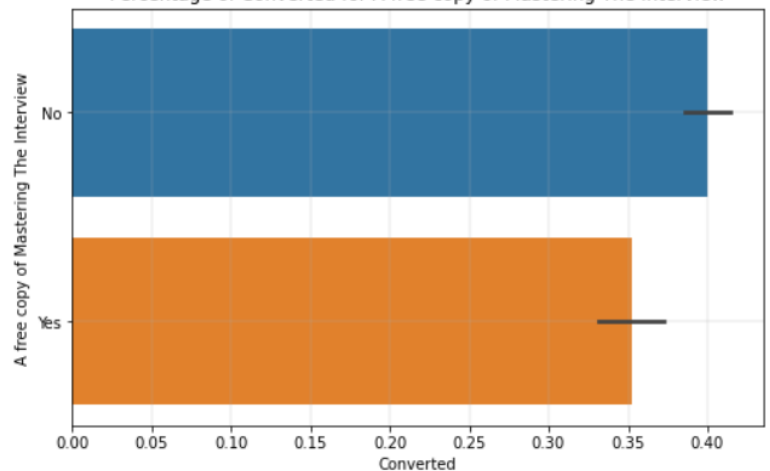
```
] : 0    0.614564  
    1    0.385436  
    Name: Converted, dtype: float64
```

- We can see that the customer who is interested in 'Mastering the interview', they are likely to be converted.

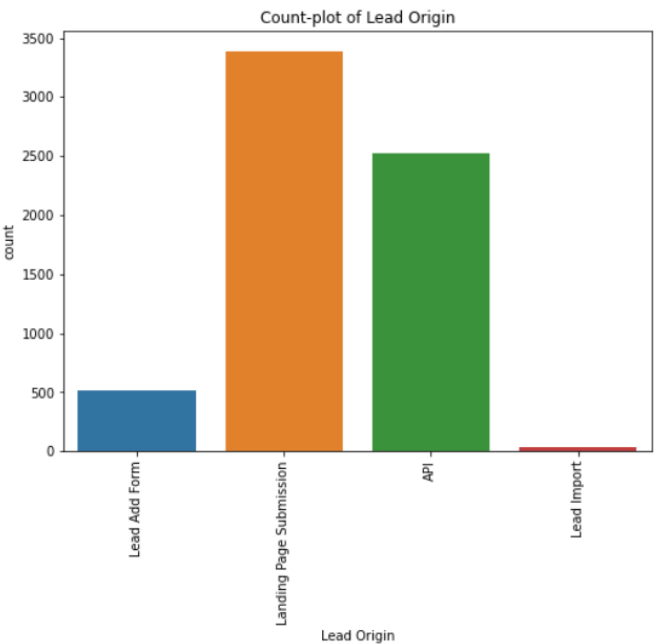
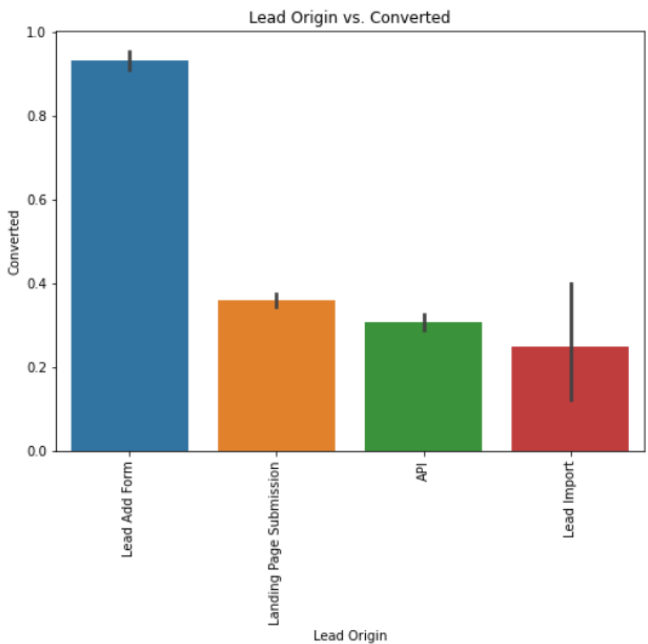
Distribution of A free copy of Mastering The Interview



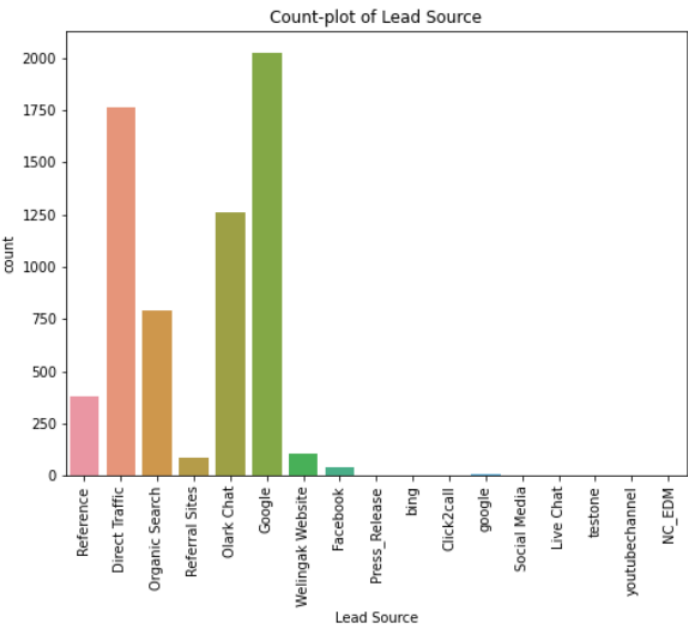
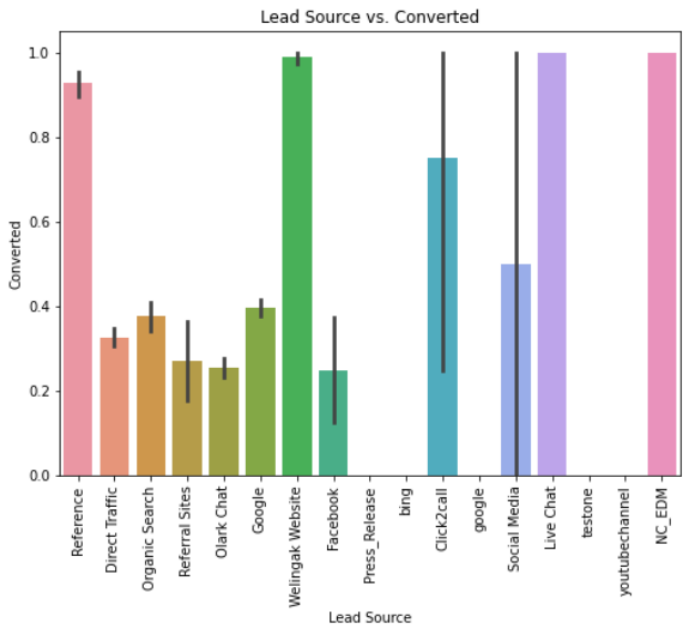
Percentage of Converted for A free copy of Mastering The Interview



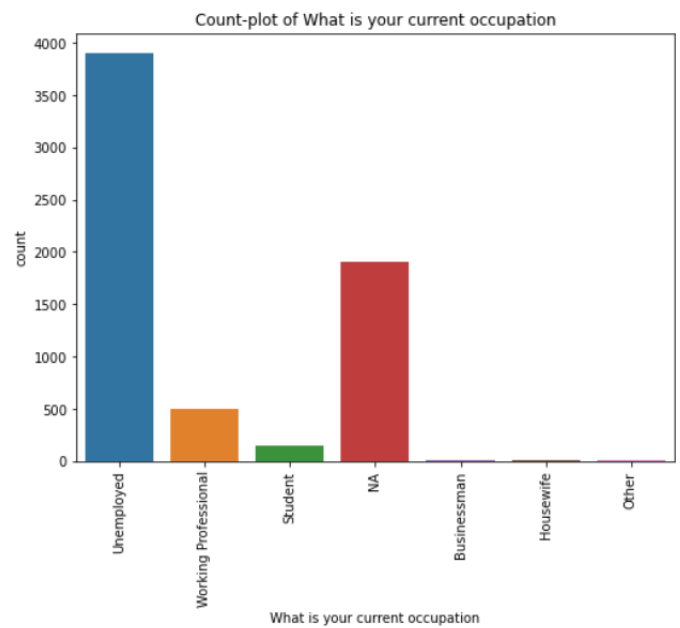
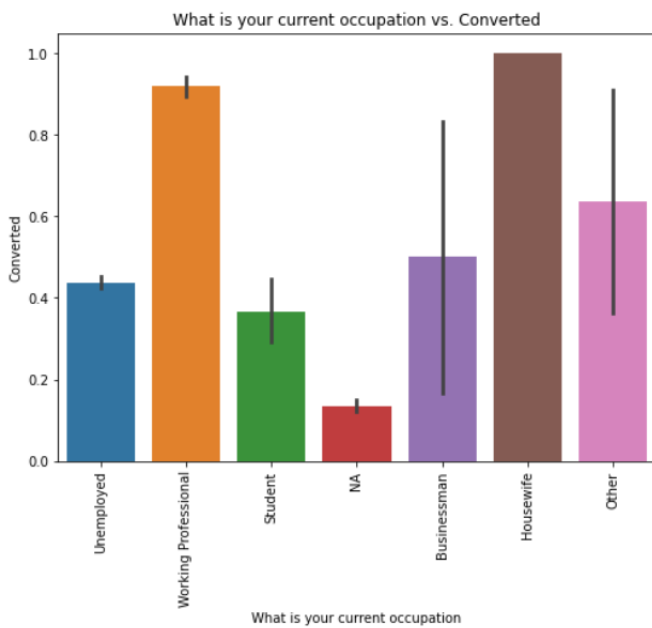
- Customers whose Lead Origin is 'Lead Add Form' are highly likely to be converted. There are approximately 500 'Lead Add Form' records in train dataset which is a good number.



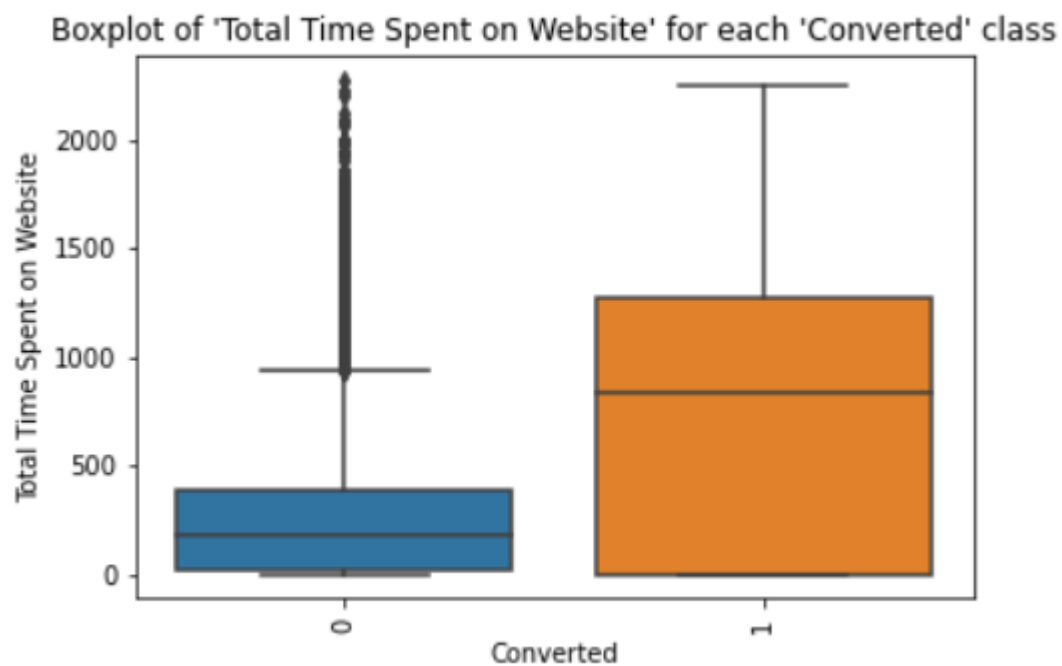
- Similarly, Customers whose Lead Source is 'Reference' are highly likely to be converted.



- Targeting working professional would be a good idea as their converted rate is very high.

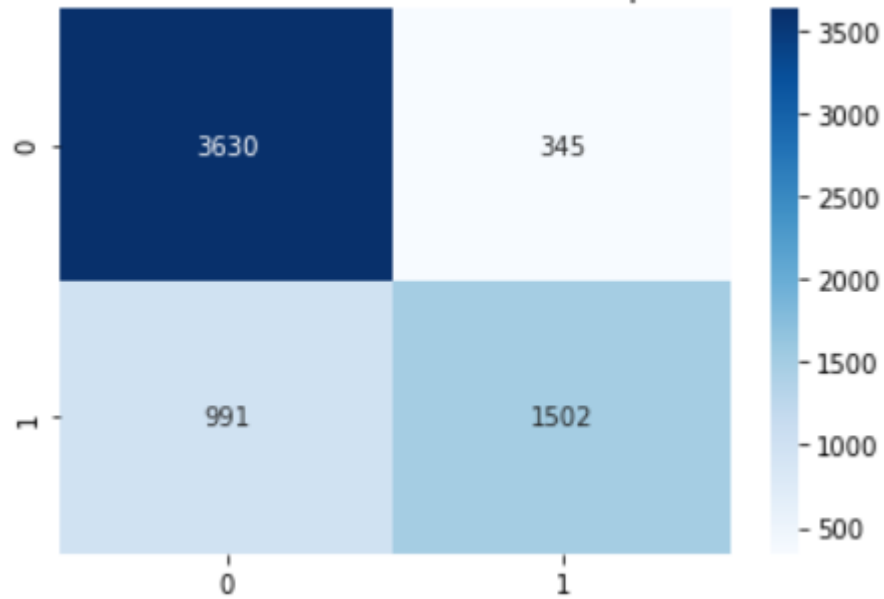


- People who are spending more time on the website are likely to be Converted.

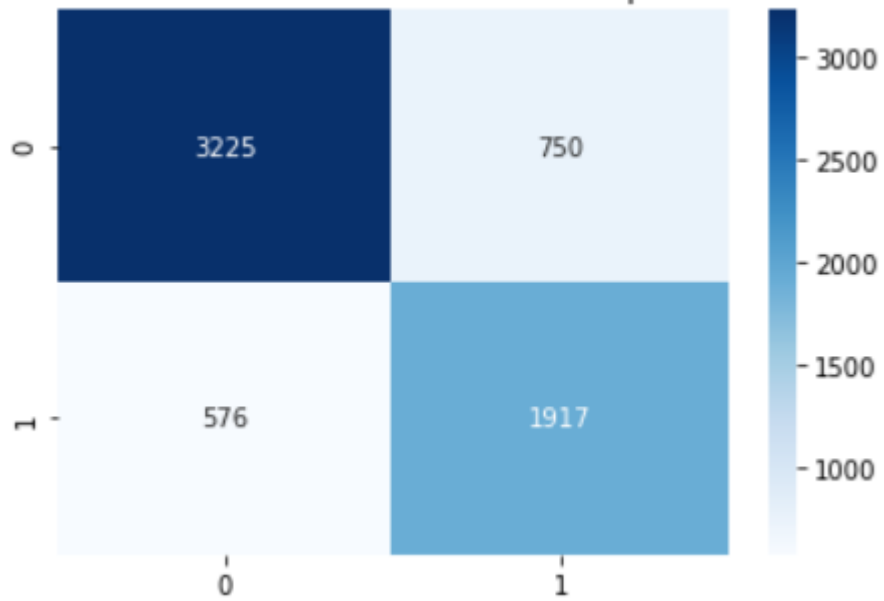


- We can see how sensitivity is increasing by decreasing the cutoff point.

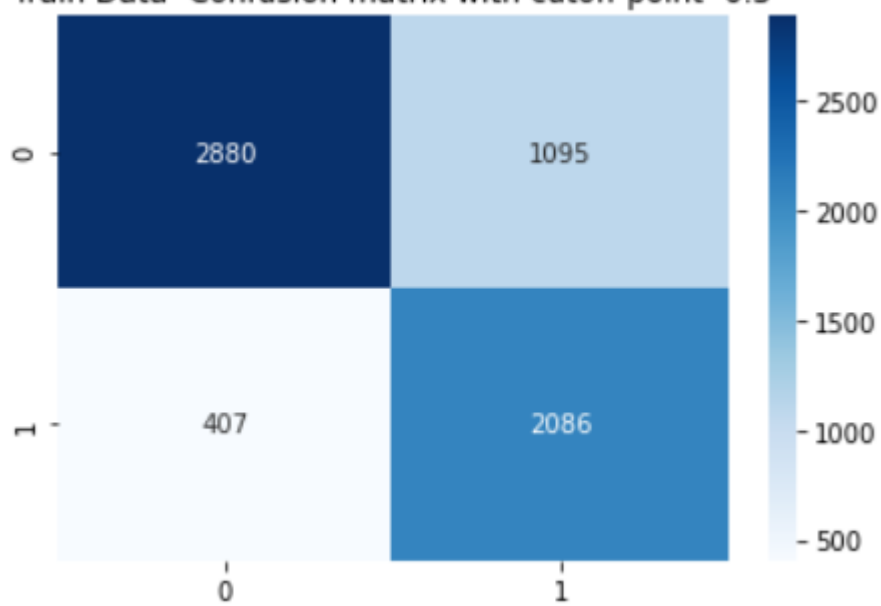
Train Data- Confusion matrix with cutoff point- 0.6



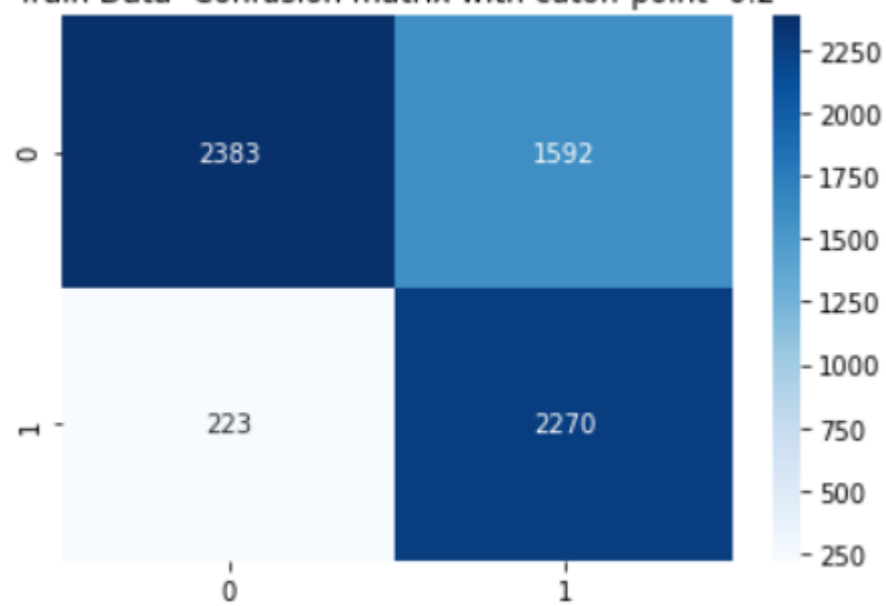
Train Data- Confusion matrix with cutoff point- 0.4

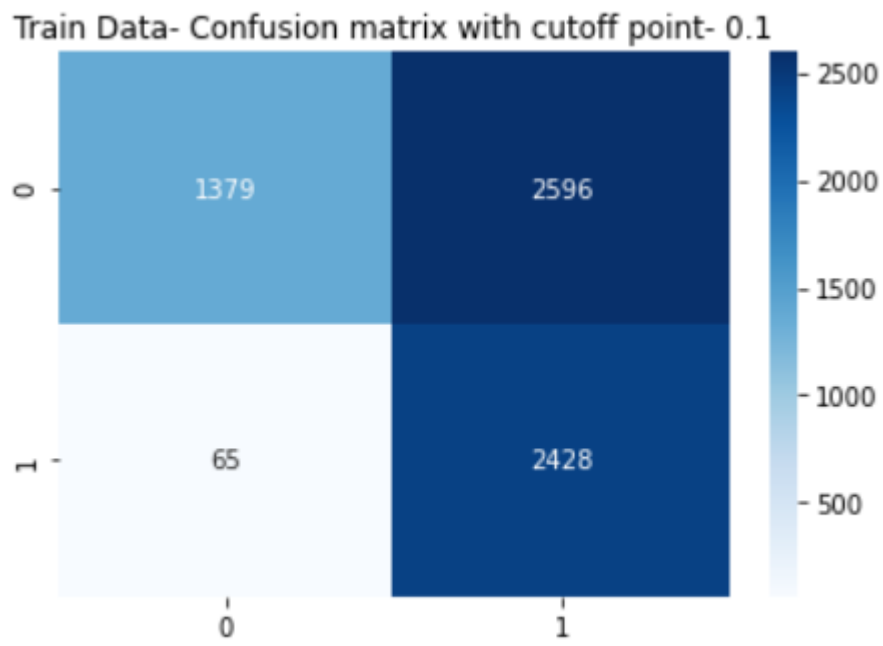


Train Data- Confusion matrix with cutoff point- 0.3



Train Data- Confusion matrix with cutoff point- 0.2





Hence, if we have to make calls aggressively then we can decrease the cutoff point, and if we want to reduce the number of calls then we can increase the cutoff point.